An Information-Theoretic Approach to Quantum Theory, II: The Formal Rules of Quantum Theory

Philip Goyal*
Cavendish Laboratory
University of Cambridge

In a companion paper [1] (hereafter referred to as Paper I), we have presented an attempt to derive the finite-dimensional abstract quantum formalism from a set of physically comprehensible assumptions. In this paper, we formulate a correspondence principle, the *Average-Value Correspondence Principle*, that allows relations between measurement outcomes which are known to hold in a classical model of a system to be systematically taken over into the quantum model of the system. Using this principle, we derive the explicit form of the temporal evolution operator (thereby completing the derivation of the abstract quantum formalism begun in Paper I), and derive many of the formal rules (such as operator rules, commutation relations, and Dirac's Poisson bracket rule) that are needed to apply the abstract quantum formalism to model particular physical systems.

I. INTRODUCTION

In order to obtain a quantum mechanical model for a particular physical system such as a particle moving in space, it is necessary to supplement the abstract quantum formalism with formal rules which explicitly determine the form of the operators that represent particular measurements performed on the system, or that represent particular symmetry transformations (such as displacement or rotation) of the frame of reference in which the system is being observed. These formal rules usually suppose that measurements are described in the framework of classical physics, so that one speaks of "a measurement of observable (or property) A" and of an operator that represents such a measurement. These formal rules can be usefully classified as follows:

(i) *Operator Rules.* Rules for writing down operators representing measurements of observables that are known functions of other, elementary, observables whose operators are given [27]. For example, such rules are needed to be able to write down the operator that represents a measurement of H, given the classical relation $H = p_x^2/2m + V(x)$, in terms of the operators x and p_x which represent measurements of x and p_x, respectively.

(ii) *Commutation Relations.* Commutation relations between operators that represent measurements of fundamental observables such as position, momentum, and components of angular momentum. The commutation relations $[\mathsf{x}, \mathsf{p}_x] = i\hbar$ and $[\mathsf{L}_x, \mathsf{L}_y] = i\hbar\mathsf{L}_z$, are the obvious examples, while Dirac's Poisson bracket rule, $[\mathsf{A}, \mathsf{B}] = i\hbar\widehat{\{A, B\}}$, is the more general rule for evaluating commutation relations, where $\{A, B\}$ is the classical Poisson bracket for observables A and B of a physical system, and $\widehat{\{A, B\}}$, A, and B are the respective operators.

(iii) *Transformation Operators.* Explicit forms of the operators that represent symmetry transformations of a frame of reference, such as the x-displacement operator $\mathsf{D}_x = -i\, d/dx$.

(iv) *Measurement–Transformation Relations.* The relations between measurement operators and transformation operators. For example, the x-displacement operator, D_x, stands in the relation $\mathsf{D}_x = \mathsf{p}_x/\hbar$ to the x-momentum measurement operator, p_x.

The physical origin of many of the above-mentioned rules is obscure. For example, although one can give simple physical arguments [28] for the operator rule which states that, if a measurement of A is represented by operator A, then a measurement of a function, $f(A)$, of A is represented by operator $f(\mathsf{A})$, the generalization of such arguments to measurements of functions of two or more observables encounters severe difficulties due to the possible non-commutativity of the operators that represent these observables. As a result of such difficulties, operator rules tend to be heuristic and tend to lead to inconsistencies when applied to particular examples [29]. Similarly, the commutation relationship $[\mathsf{x}, \mathsf{p}_x] = i\hbar$ is typically obtained from Schroedinger's equation or from Dirac's Poisson bracket rule, both of whose derivations involve abstract assumptions whose physical origin is obscure.

Recent work on the elucidation of the physical origin of the quantum formalism either focuses exclusively on the derivation of the abstract quantum formalism [2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13] or is concerned with the derivation of the Schroedinger equation directly from informational ideas without taking the abstract quantum formalism as a given [14, 15, 16]. Consequently, the question of what additional physical ideas are needed to obtain the formal rules described above *given* the abstract quantum formalism has received relatively little recent attention.

Operator rules have been discussed in a few publications, for example in [17, 18], but a derivation of the operator rules on the basis of a physical principle, taking the

*Electronic address: pg247@cam.ac.uk

abstract formalism as a given, has not been successfully completed. The most recent systematic attempt to derive the commonly employed commutation relations and measurement–transformation relations in such a manner is found in [19] (see also Refs. [20, 21]). In [19], the commutation relationships for the operators that represent the Galilean group of transformations are first derived by exploiting the group structure of the classically-described transformations. Then, by establishing the relation between these transformation operators and particular measurement operators, commutation relations for these measurement operators are obtained. However, this approach implicitly makes use of operator rules, and relies upon auxiliary assumptions, such as the assumption that certain measurement operators are unchanged by the action of particular symmetry operators, whose physical origin is unclear.

In this paper, we show that, starting from the abstract quantum formalism, it is possible to derive the above-mentioned formal rules in a straightforward manner from a correspondence principle. Roughly speaking, this principle, the *Average-Value Correspondence Principle* (AVCP), asserts that, in a classical experiment, if a relation holds between the outcomes of a set of measurements performed on a physical system, then the same relation holds *on average* in a corresponding, suitably-defined quantum experiment on the same physical system, the average being taken over infinitely many trials of the quantum experiment.

This paper is organized as follows. We begin in Sec. II by formulating the AVCP. In Sec. III, the AVCP is used to obtain several generalized operator rules which connect the average values of operators at different times, from which the commonly used operator rules of quantum theory follow as a special case. Using the AVCP and Postulate 3.4 (from Paper I), we then derive the explicit form of the temporal evolution operator, which completes the derivation of the finite-dimensional abstract quantum formalism begun in Paper I.

Next, in Sec. IV, taking the infinite-dimensional form of the abstract quantum formalism as a given, we use the AVCP to derive many of the commonly employed formal rules of quantum theory, namely (a) the commutation relations $[\mathsf{x}, \mathsf{p}_x] = i\hbar$ and $[\mathsf{L}_x, \mathsf{L}_y] = i\hbar\mathsf{L}_z$, and Dirac's Poisson bracket rule, (b) the explicit form of the operators for displacements and rotations, and (c) the relation between momentum and displacement operators, and between angular momentum and rotation operators.

Finally, in Sec. V, we show that, in the functions $f(\chi_i) = \pm\cos(a\chi_i + b)$ and $\tilde{f}(\chi_i) = \pm\sin(a\chi_i + b)$ derived in Paper I, the signs can be taken to be positive and the constants a and b can be taken to have the values $a = 1$ and $b = 0$ without loss of generality.

We note that the treatment of the formal rules is illustrative rather than exhaustive, so that only the most commonly encountered measurement and transformation operators which are needed to formulate non-relativistic and relativistic quantum mechanics have been discussed. Many other formal rules (such as the operators for Galilei transformations, temporal displacement, and discrete transformations such as spatial inversion) can be obtained by arguments which closely follow those presented. The paper concludes with a discussion of the results obtained.

II. THE AVERAGE-VALUE CORRESPONDENCE PRINCIPLE

A. Introduction

Suppose that, as described in Paper I, a quantum model $\mathbf{q}(N)$, of dimension N, has been constructed to describe an abstract experimental set-up consisting of a source of identical systems, a measurement set $\mathcal{A}$, and an interaction set $\mathcal{I}$. The measurements in $\mathcal{A}$, and the degenerate forms of measurements in $\mathcal{A}$ [30], are represented by N-dimensional Hermitian operators (possibly with degenerate eigenvalues), and shall be called *quantum* measurements.

Suppose that quantum measurement $\mathbf{A}$, with operator A, represents a measurement that is classically described as a measurement of some observable A (which we shall henceforth abbreviate to "quantum measurement $\mathbf{A}$ (or operator A) represents a measurement of A"). Suppose that we wish to determine whether there is a quantum measurement that represents a measurement of, say, A^2 and, if so, to determine the operator which can be said to represent this measurement. One can imagine that, classically, a measurement of A^2 is implemented by a process where a measurement of A is performed and the outcome is then squared. If this implementation is described in the quantum model, it follows that, if the input state is an element, v_i, of an orthonormal set of eigenvectors of A, the output state of the process is v_i and the observed result of the process is a_i^2, where $\mathsf{A}\mathsf{v}_i = a_i\mathsf{v}_i$ $(i = 1, 2, \ldots, N)$. According to this line of reasoning, it follows at once that the operator A^2 therefore represents a measurement of A^2.

However, this simple argument does not readily generalize. For example, suppose that quantum measurements $\mathbf{A}$ and $\mathbf{B}$, with operators A and B, represent measurements of A and B, respectively. Suppose that we wish to determine the operator (if any exists) that represents a measurement of $A + B$. Classically, one imagines that such a measurement is implemented by making measurements of A and B simultaneously on a system, and adding the outcomes. However, if $[\mathsf{A}, \mathsf{B}] \neq 0$, this implementation cannot be described without modification in the quantum framework since the measurements cannot be performed at the same time, and the order in which they are performed is of potential significance.

Now, the conventional operator rules of quantum theory assert that a measurement of $A+B$ is represented by the operator $\mathsf{A}+\mathsf{B}$. Although this seems entirely reasonable on a formal, symbolic level, it is not clear in what

physical sense $\mathsf{A}+\mathsf{B}$ can be said to 'represent' the measurement since, whereas the eigenvectors and eigenvalues of A directly reflect the output states and outcome values obtained when measurement $\mathbf{A}$ is performed on the system, the eigenvectors of $\mathsf{A}+\mathsf{B}$ do not, in general, coincide with the eigenvectors of either A or B, and the eigenvalues of $\mathsf{A}+\mathsf{B}$ do not, in general, coincide with the possible results of any plausible implementation (described in the quantum model) of a measurement of $A+B$ that involves performing measurements of A and B and combining their outcomes. For example, a measurement represented by operator $\mathsf{S}_x+\mathsf{S}_z$ on a spin-1/2 particle has possible outcomes $\pm\hbar/\sqrt{2}$, whereas the possible results of an implementation where a measurement of S_x is followed by a measurement of S_z (or vice-versa), and their outcomes are added, are $\pm\hbar, 0$.

The above observations illustrate the difficulty of obtaining a physical understanding of the operator rules of quantum theory even in simple cases of interest. Below, we shall formulate a physical principle which gives a clear physical meaning to the sense in which an operator can be said to represent a classically-described measurement, and, in the majority of cases of interest, uniquely determines the operator which represents such a measurement.

1. *Implementations of classically-described measurements*

Consider again a measurement that, from the classical standpoint, is said to be a measurement of A^2, where A is some observable. Classically, one can imagine a measurement of A^2 being implemented in one of three different ways (see Fig. 1): (i) make a measurement of A on one copy of the system, and square the outcome; (ii) make two immediately successive measurements of A on one copy of the system, and multiply the two outcomes; or (iii) make two simultaneous measurements of A on two separate copies of the system prepared in the same state, and multiply the two outcomes. Although the first of these implementations is the one we have considered above, all three implementations yield the same result when modeled classically, and so can be regarded as equally valid implementations of a measurement of A^2.

Now, perhaps surprisingly, when described using the quantum model, these implementations do not, in general, yield the same expected results. Consider a quantum experimental arrangement employing the first implementation. In each run of the experiment, one copy of the system is prepared in some given state. Let the probability that a measurement of A yields outcome i ($i=1,2,\ldots,N$), with value a_i, be denoted P_i. Then, the expected result is given by

$$\begin{aligned}\text{Expected result} &= \sum_i (a_i)^2 P_i \\ &= \overline{a^2}\end{aligned} \tag{1}$$

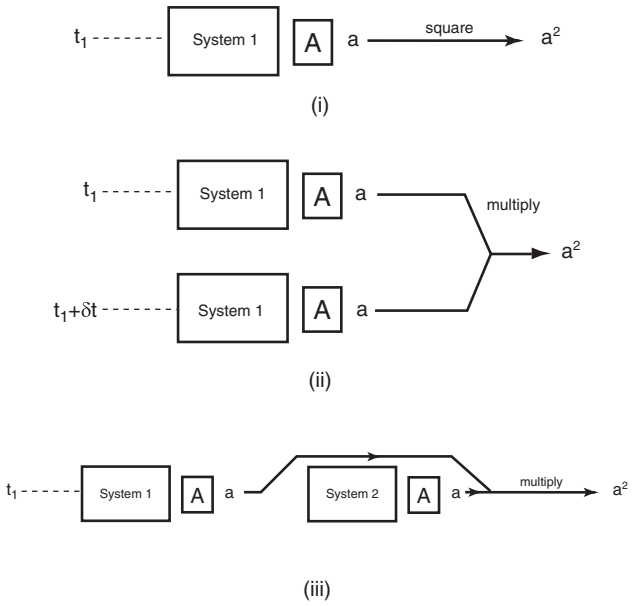


FIG. 1: Three implementations of the measurement classically described as "a measurement of A^2". In (i), a measurement of A is made on one copy of the system, and the outcome is squared to give the result; in (ii), two immediately successive measurements of A are made on one copy of the system, and the result is obtained by multiplying the two outcomes; and (iii) two simultaneous measurements of A are made on two separate copies of the system prepared in the same state, and the result is obtained by multiplying the two outcomes. In a classical model of this situation, each implementation yields the same result. However, in the quantum model of these implementations, the expected result of (i) and (ii) is $\overline{a^2}$ whereas the expected result of (iii) is $(\overline{a})^2$.

One finds that implementation (ii) yields the same expected result. However, in implementation (iii), where, in each run of the experiment, two copies of the system are prepared in the same state, one obtains

$$\begin{aligned}\text{Expected result} &= \sum_{i,j} (a_i a_j) P_i P'_j \\ &= (\overline{a})^2,\end{aligned} \tag{2}$$

with P_i and P'_j denoting respectively the probabilities that the measurements of A on the first and second copy yield outcomes i and j ($i,j=1,2,\ldots,N$).

In this example, the differences between the implementations as viewed in the quantum model arise due to the fact that, in the quantum framework, an immediate repetition of a measurement on the same copy of a system is different from performing a second simultaneous measurement on a separate, identically-prepared, copy of the system.

In the case of a measurement that is classically described as a measurement of $A+B$, where measurements of A and B are assumed to occur at the same time, one needs to take into account the additional fact that, in the quantum framework, the order in which two mea-

surements is performed is of possible importance. Accordingly, one can imagine implementing a measurement of $A+B$ in one of at least three different ways (see Fig. 2): (i) make a measurement of A, and then a measurement of B, on one copy of the system, and add the outcomes; (ii) make a measurement of B, and then a measurement of A, on one copy of the system, and add the outcomes; (iii) make a measurement of A on one copy of the system and, simultaneously, a measurement of B on a second copy of the system prepared in the same state as the first copy, and add the outcomes. Once again, in a classical model of this situation, the results agree. However, if one calculates the expected results in the quantum model, one finds that, if $[\mathsf{A}, \mathsf{B}] \neq 0$, all three will, in general, disagree with one another.

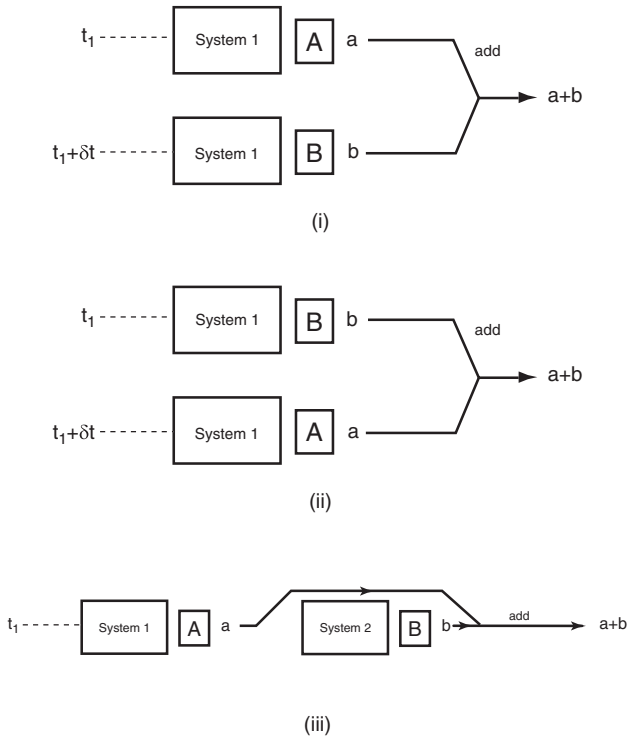


FIG. 2: Three implementations of the measurement classically described as "a measurement of $A + B$". In (i), a measurement of A, and then a measurement of B, is made on one copy of the system, and the outcomes are then added to give the result; in (ii), a measurement of B, and then a measurement of A, is made on one copy of the system, and the outcomes are then added to give the result; and (iii) simultaneous measurements of A and B are made on two separate copies of the system prepared in the same state, and the result is obtained by adding the two outcomes. In a classical model of this situation, each implementation yields the same result. However, in the quantum model of this situation, the three implementations do not, in general, yield the same expected results.

As these examples illustrate, different implementations of the same classically-described measurement do not, in general, give the same expected result in the quantum model of the experimental arrangement. Thus, whereas a classical description of a measurement, such as "a measurement of x^2", adequately describes the intended measurement insofar as the outcome is concerned in the classical framework, such a description allows for more than one implementation which, in general, do not yield the same expected result when modeled in the quantum framework.

2. The average-value condition

The existence of different implementations of the same classically-described measurement immediately raises two questions. First, are these implementations, in some sense, equally valid in the quantum framework, or it is possible to find some reasonable physical basis upon which to select particular implementations and regard these as more fundamental than the others? Second, is it possible to find operators that represent the selected implementations, and, if so, do all the selected implementations of a given measurement have the same operator representation?

To answer these questions, we begin by observing that, although the above implementations are all regarded as *bone fide* measurements in the classical model, a measurement is only describable as such in the quantum model, and so can be called a quantum measurement, if it can be represented by a Hermitian operator which represents a single measurement performed upon one copy of the system at a particular time. So, for example, although implementation (iii) of a measurement of A^2 can be modeled in the quantum framework, the process *as a whole* cannot be described as a quantum measurement since it involves two separate measurements. In contrast, implementation (i) of a measurement of A^2 can be described as a quantum measurement since it only involves a single measurement on one copy of the system.

However, although an implementation modeled in the quantum framework that involves more than one measurement cannot itself be regarded as a quantum measurement, we can reasonably ask whether it is possible to find a quantum measurement, $\mathbf{C}$, with operator C, which, in some sense to be determined, can nonetheless be said to *represent* the implementation.

At the outset, we note that it makes no sense to require that measurement $\mathbf{C}$ yield the same outcome as a given implementation since measurement outcomes are only probabilistically determined in the quantum framework. However, we *can* impose the simple condition that, over an infinite number of runs, the average value of measurement $\mathbf{C}$ should coincide with the average result obtained from the implementation for any initial state of the system.

For example, in the case of an implementation of a measurement of A^2, measurement $\mathbf{C}$ which, by hypothesis, represents the implementation, must be such that, for all states of the system, $\langle \mathsf{C} \rangle$ is equal to the expected result obtained from the implementation. In the case of imple-

mentations (i) and (ii) described above, using Eq. (1), we accordingly obtain the condition that the relation

$$\begin{aligned}\langle \mathsf{C}\rangle &= \overline{a^2} \\ &= \langle \mathsf{A}^2\rangle\end{aligned} \tag{3}$$

must hold for all states, v, of the system. From this condition, we can conclude that

$$\mathsf{C} = \mathsf{A}^2. \tag{4}$$

In the case of implementation (iii), using Eq. (2), we obtain the condition that the relation

$$\begin{aligned}\langle \mathsf{C}\rangle &= (\overline{a})^2 \\ &= \langle \mathsf{A}\rangle^2\end{aligned} \tag{5}$$

must hold for all v. By diagonalizing A, one can readily show that this condition implies that A is a multiple of the identity, which represents a trivial measurement that yields the same outcome irrespective of the state of the system. Therefore, implementation (iii) does not satisfy the above average-value condition in the case of any non-trivial measurement of A, and can therefore be reasonably eliminated. Hence, in this case, the average-value condition is sufficiently strong so as to be able to pick out implementations (i) and (ii), and, since the average-value condition also implies that these implementations are both represented by the operator, A^2, one can unambiguously conclude that a measurement of A^2 is represented by the operator A^2.

Proceeding in a similar way, restricting ourselves for the time being to measurements $\mathbf{A}$ and $\mathbf{B}$ that are not sub-system measurements [31], one finds that, in the case of a measurement of $C = A+B$, only implementation (iii) is possible if $[\mathsf{A}, \mathsf{B}] \neq 0$, which then yields the operator

$$\mathsf{C} = \mathsf{A} + \mathsf{B}. \tag{6}$$

If $[\mathsf{A}, \mathsf{B}] = 0$, then all three implementations are possible, and all yield the same operator, C, as above.

Finally, in the case of a measurement of AB, with $[\mathsf{A}, \mathsf{B}] = 0$, one finds that the implementation must the one where measurements $\mathbf{A}$ and $\mathbf{B}$ are performed on the same copy of the system, in which case the operator AB is obtained.

Hence, we see that the average-value condition is sufficient to yield a unique operator representation for the measurements considered above. Based on the above considerations, we can tentatively formulate the following general rule: in the case of a measurement which has an implementation that contains two elementary measurements (not sub-system measurements) represented by commuting operators, it is possible to find a quantum measurement that represents the implementation if the two elementary measurements are performed on the same copy of the system; but, when the operators do not commute, the elementary measurements must be performed on different copies of the system.

We now consider implementations of a measurement of AB in the case when $[\mathsf{A}, \mathsf{B}] \neq 0$. The general rule just given suggests that we should consider the implementation of this measurement where the measurements of A and B are performed on different copies of the system. This implementation has the expected value

$$\sum_i \sum_j (a_i b_j) P_i P'_j = \langle \mathsf{A}\rangle\langle \mathsf{B}\rangle, \tag{7}$$

with P_i and P'_j denoting respectively the probabilities that the measurement of A on the first copy and the measurement of B on the second copy yield outcomes i and j $(i, j = 1, 2, \ldots, N)$, and a_i, b_j respectively denoting the values of the ith and jth outcomes of measurements $\mathbf{A}$ and $\mathbf{B}$. Imposing the above average-value condition, the operator C that represents this implementation must satisfy the condition

$$\langle \mathsf{C}\rangle = \langle \mathsf{A}\rangle\langle \mathsf{B}\rangle \tag{8}$$

for all v. However, for non-commuting A and B, one finds that C cannot be found such that this relation is satisfied for all v. We note, however, that with

$$\mathsf{C} = \frac{1}{2}(\mathsf{AB} + \mathsf{BA}), \tag{9}$$

equation (8) holds for the eigenstates of A and the eigenstates of B, which suggests that we may be able to weaken the average-value condition in this case so that we only require that Eq. (8) holds for these eigenstates. However, as we shall illustrate later (Sec. III A 1), the application of Eq. (9) can lead to inconsistencies. Consequently, we conclude that, from the point of view of the average-value condition, it is not possible to find an operator that represents a measurement of AB when the operators A and B do not commute. More generally, we find that, when $[\mathsf{A}, \mathsf{B}] \neq 0$, it is necessary to exclude measurements of $f(A, B)$, with f analytic (so that f has a well-defined polynomial expansion in A and B), where the polynomial expansion of f contains product terms.

Finally, in the case of a measurement which has a classical implementation which contains two elementary measurements that are performed on different sub-systems of a composite system, one finds that the implementation can be represented by a quantum measurement irrespective of whether the two measurements are performed on the same or on different copies of the system, and that the different possible implementations are represented by the same operator.

3. Generalizations

In the examples above, we have considered implementations of a classically-described measurement in which the measurement and the elementary measurements in the implementation are performed at the same time, and

in the same frame of reference. However, as the following examples illustrate, these are unnecessary restrictions.

First, classically, for a non-relativistic particle of mass m, one can implement a measurement of x performed at time $t + \delta t$ by a process where measurements of x and p_x are performed at time t, and the function $x + p_x\,\delta t/m$ is then computed.

Second, one can implement a measurement of x' on a particle in the reference frame S' that is displaced along the x-axis relative to frame S by performing a measurement of x in frame S, and computing the appropriate function $x' = f(x)$ that relates x and x'.

The above considerations regarding the implementation of a classically-described measurement are applicable without change to the case where the measurement and the elementary measurements in its implementation are performed at different times or in different frames of reference. Below, we shall accordingly generalize our notion of the implementation of a classically-described measurement.

B. Statement of the Principle

We shall now state a general principle which incorporates the observations made above concerning the average-value condition. An illustrative example is given in Fig. 3.

Average-Value Correspondence Principle Consider a classical idealized experiment in which a system (possibly a composite system) is prepared in some state at time t_0, and is allowed to evolve in a given background. Suppose that a measurement of $A^{(m)}$ ($m \geq 2$), performed on the system at time t_2 with outcome $a^{(m)}$, has a classical implementation in which measurements of $A^{(1)}, A^{(2)}, \ldots, A^{(m-1)}$ are performed upon one copy of the system at time t_1, and their respective outcomes, denoted $a^{(1)}, a^{(2)}, \ldots, a^{(m-1)}$, are then used to compute the result $f(a^{(1)}, a^{(2)}, \ldots, a^{(m-1)})$, where f is an analytic function, so that the relation

$$a^{(m)} = f(a^{(1)}, a^{(2)}, \ldots, a^{(m-1)}) \qquad (*)$$

holds for all initial (classical) states of the system.

Consider the case where the quantum measurements $\mathbf{A}^{(1)}, \mathbf{A}^{(2)}, \ldots, \mathbf{A}^{(m)}$, with operators $\mathsf{A}^{(1)}, \mathsf{A}^{(2)}, \ldots, \mathsf{A}^{(m)}$, represent the measurements of $A^{(1)}, A^{(2)}, \ldots, A^{(m)}$, respectively. Then, consider the following idealized quantum experimental arrangement consisting of several set-ups, each consisting of identical sources and backgrounds, where, in each set-up, a copy of the system is prepared in the same initial state, v_0, at time t_0.

In one set-up, only measurement $\mathbf{A}^{(m)}$ is performed (at time t_2) and, for any i, j with $i \neq j$ and $i, j \leq m-1$, the measurements $\mathbf{A}^{(i)}, \mathbf{A}^{(j)}$ are performed (at time t_1) in the same set-up in the quantum experiment if $\left[\mathsf{A}^{(i)}, \mathsf{A}^{(j)}\right] = 0$ provided that, if the system is composite, the measurements are performed on the same sub-system; and are performed (at time t_1) in different set-ups if $\left[\mathsf{A}^{(i)}, \mathsf{A}^{(j)}\right] \neq 0$.

Let the outcomes of the measurements $\mathbf{A}^{(1)}, \ldots, \mathbf{A}^{(m)}$ in any given run of the experimental arrangement be denoted $a^{(1)}, \ldots, a^{(m)}$, respectively. The function $f(a^{(1)}, a^{(2)}, \ldots, a^{(m-1)})$ is defined as *simple* provided that its polynomial expansion contains no terms involving a product of eigenvalues belonging to measurements whose operators do not commute. If f is simple, then $(*)$ holds on average, the average being taken over an infinitely large number of runs of the experiment.

The above principle can be generalized in a number of ways, for example to the case where the measurements of $A^{(1)}, \ldots, A^{(m-1)}$ are not performed at the same time. However, these generalizations are unnecessary for the derivations of the usual formal rules of quantum theory, and are therefore not discussed here.

III. GENERALIZED OPERATOR RULES AND THE TEMPORAL EVOLUTION OPERATOR

A. Generalized Operator Rules

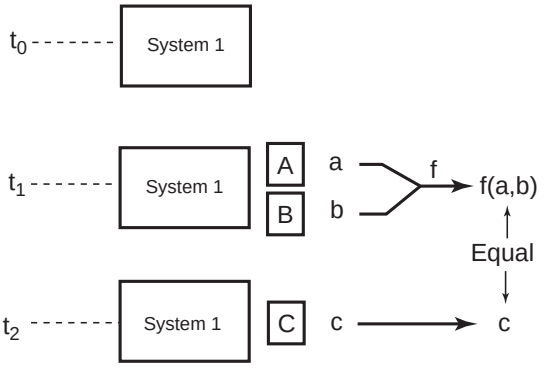
We will now apply the AVCP to derive operator relations which hold when the function f takes various useful forms. In each instance of f, we shall first derive a generalized operator rule which relates the expected values of the relevant operators at *different times*. Then, taking the special case when the expected values are computed at the same time, we obtain the corresponding operator rule which relates the operators themselves.

We shall consider a classical experiment where a system is subject to measurements of A and B at time t_1, and to a measurement of C at time t_2. We shall suppose that measurement of C, with outcome c, can be implemented by an arrangement in which the measurements of A and B are performed, with respective outcomes a and b, and the function $f(a, b)$ then computed, so that the relation

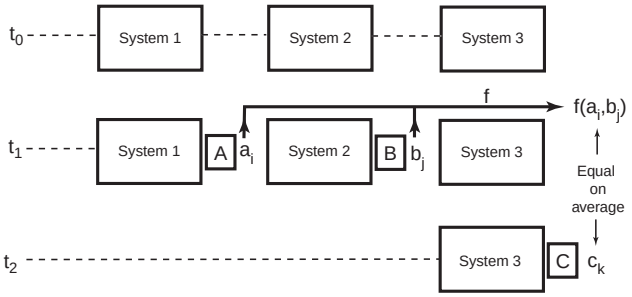
$$c = f(a, b) \qquad (10)$$

holds for all initial states of the system.

In a quantum model of the appropriate experimental arrangement, let the operators that represent these measurements be denoted A, B, and C, respectively. To simplify the presentation, we shall only consider the case where these operators have finite dimension, N; the results obtained below can be readily shown to apply in the infinite dimensional case. Let the elements of orthonormal sets of eigenvectors of A, B and C be denoted $\mathsf{v}_i, \mathsf{v}'_j$,



(i)



(ii)

FIG. 3: An example of the application of the AVCP. (i) A classical experiment showing the measurements of A, B and C performed at times t_A, t_B, and t_C, respectively, with outcomes denoted as a, b, and c, respectively. Here, $t_A = t_B = t_1$ and $t_C = t_2$. Suppose that one finds that the relation $c = f(a, b)$ holds for all initial states of the system. (ii) The corresponding quantum experiment. Three copies of the system are prepared in the same initial state, v_0, at time t_0, and are placed in identical backgrounds. In this example, it is assumed that the operators A and B do not commute. Hence, by the AVCP, measurements **A** and **B** are performed on different copies of the system. Measurement **C** is performed on a separate copy of the system. In any given run of the experiment, the probabilities that measurements **A**, **B** and **C** yield outcome values a_i, b_j and c_k $(i, j, k = 1, \ldots, N)$, are P_i, P'_j and P''_k, respectively. The AVCP then asserts that, provided the polynomial expansion of $f(a, b)$ contains no product terms involving a and b, the relation $\overline{c} = \overline{f(a,b)}$ holds for all initial states, v_0, of the system, where the average is taken over an infinite number of runs of the experiment; that is, $\sum_k c_k P''_k = \sum_{ij} f(a_i, b_j) P_i P'_j$ for all v_0.

and v''_k, respectively $(i, j, k = 1, 2, \ldots, N)$, let the corresponding eigenvalues be denoted a_i, b_j and c_k, and let the probabilities of the ith, jth and kth outcomes of measurements **A**, **B** and **C** in any given experimental arrangement be denoted by P_i, P'_j and P''_k, respectively.

Case 1. f is a function of a only.

In this case, the quantum experiment simply consists of two identical set-ups, involving two copies of the system, where **A** is performed on one copy at time t_1 and **C** on the other copy at time t_2. Since function f is simple, by the AVCP, the relation

$$\sum_k c_k P''_k = \sum_i f(a_i) P_i \tag{11}$$

holds for all initial states, v_0, of the system. Explicitly,

$$\begin{aligned} P_i &= \left|\mathsf{v}_i^\dagger \mathsf{v}_{t_1}\right|^2 \\ P''_k &= \left|\mathsf{v}''^\dagger_k \mathsf{v}_{t_2}\right|^2, \end{aligned} \tag{12}$$

with v_t being the state of the relevant copy of the system at time t. Hence, we can write Eq. (11) as

$$\mathsf{v}^\dagger_{t_2}\left(\sum_k \mathsf{v}''_k \mathsf{v}''^\dagger_k c_k\right)\mathsf{v}_{t_2} = \mathsf{v}^\dagger_{t_1}\left(\sum_i \mathsf{v}_i \mathsf{v}_i^\dagger f(a_i)\right)\mathsf{v}_{t_1}. \tag{13}$$

Noting that

$$\begin{aligned} f(\mathsf{A}) &= \sum_i \mathsf{v}_i \mathsf{v}_i^\dagger f(a_i) \\ \mathsf{C} &= \sum_k \mathsf{v}''_k \mathsf{v}''^\dagger_k c_k, \end{aligned} \tag{14}$$

we obtain the relation

$$\langle \mathsf{C} \rangle_{t_2} = \left\langle f(\mathsf{A}) \right\rangle_{t_1}, \tag{15}$$

which holds for all v_0. We can summarize the above result in the form of the *generalized function rule*:

$$c(t_2) = f\left(a(t_1)\right) \mapsto \langle \mathsf{C} \rangle_{t_2} = \left\langle f(\mathsf{A}) \right\rangle_{t_1} \quad \forall \mathsf{v}_0, \tag{16}$$

where, for clarity, the times at which the outcomes are obtained has been explicitly indicated. In the special case where $t = t_1 = t_2$, we obtain the usual operator rule, the *function rule*:

$$c = f(a) \mapsto \mathsf{C} = f(\mathsf{A}). \tag{17}$$

Case 2. $f(a, b) = f_1(a) + f_2(b)$

It is necessary to consider three sub-cases. First, if measurements **A** and **B** have commuting operators and, in the case of a composite system, if they are sub-system measurements performed on the same sub-system, then, by the AVCP, they are performed on the same copy of the system in the quantum experiment. Since f is simple, the AVCP applies, so that

$$\begin{aligned} \sum_k c_k P''_k &= \sum_i \left(f_1(a_i) + \sum_j f_2(b_j) P'_{j|i} \right) P_i \\ &= \sum_i \left(f_1(a_i) + f_2(b_i) \right) P_i, \end{aligned} \tag{18}$$

holds for all initial states, v_0, of the system. Here, the notation $P'_{j|i}$ is the probability that measurement **B** yields outcome j given that **A** has yielded outcome i; in this case, $P'_{j|i} = \delta_{ij}$. From the above relation, we obtain the generalized operator relation

$$\langle \mathsf{C} \rangle_{t_2} = \big\langle f_1(\mathsf{A}) \big\rangle_{t_1} + \big\langle f_2(\mathsf{B}) \big\rangle_{t_1}, \tag{19}$$

which holds for all initial states, v_0. In the special case where $t_1 = t_2$, we obtain the operator relation,

$$\mathsf{C} = f_1(\mathsf{A}) + f_2(\mathsf{B}). \tag{20}$$

Second, in the case where measurements **A** and **B** are sub-system measurements performed on different sub-systems, they can, by the AVCP, be performed on the same copy of the system, in which case we obtain the same results as above.

Third, if measurements **A** and **B** have non-commuting operators, then, by the AVCP, they are performed on different copies of the system in the quantum experiment. Since f is simple, the AVCP again applies, so that the relation

$$\sum_k c_k P''_k = \sum_i f_1(a_i) P_i + \sum_j f_2(b_j) P'_j \tag{21}$$

holds for all initial states of the system, which yields the same relation as in Eq. (19).

Hence, combining the foregoing three sub-cases, we obtain the *generalized sum rule*:

$$\begin{aligned} c(t_2) &= f_1\left(a(t_1)\right) + f_2\left(b(t_1)\right) \\ &\mapsto \langle \mathsf{C} \rangle_{t_2} = \big\langle f_1(\mathsf{A}) \big\rangle_{t_1} + \big\langle f_2(\mathsf{B}) \big\rangle_{t_1} \quad \forall \mathsf{v}_0. \end{aligned} \tag{22}$$

In the special case where $t = t_1 = t_2$, we obtain the *sum rule*:

$$c = f_1(a) + f_2(b) \mapsto \mathsf{C} = f_1(\mathsf{A}) + f_2(\mathsf{B}). \tag{23}$$

More generally, consider a classical experiment where measurements of $A^{(1)}, A^{(2)}, \ldots, A^{(m-1)}$ are performed on a system at time t_1 and a measurement of $A^{(m)}$ at time t_2, with outcomes $a^{(1)}, \ldots, a^{(m)}$, respectively. Suppose that the relation

$$a^{(m)} = f(a^{(1)}, a^{(2)}, \ldots, a^{(m-1)}), \tag{24}$$

where

$$f(a^{(1)}, a^{(2)}, \ldots, a^{(m-1)}) = f_1(a^{(1)}) + \cdots + f_{m-1}(a^{(m-1)}), \tag{25}$$

holds in the classical model for all initial states, and that, in the quantum model, the measurements are represented by the operators $\mathsf{A}^{(1)}, \mathsf{A}^{(2)}, \ldots, \mathsf{A}^{(m)}$, respectively. Then, one finds that the AVCP implies the generalized operator rule:

$$\begin{aligned} a^{(m)}(t_2) &= \sum_{l=1}^{m-1} f_l\big(a^{(l)}(t_1)\big) \\ &\mapsto \big\langle \mathsf{A}^{(m)} \big\rangle_{t_2} = \sum_{l=1}^{m-1} \big\langle f_l(\mathsf{A}^{(l)}) \big\rangle_{t_1} \quad \forall \mathsf{v}_0. \end{aligned} \tag{26}$$

Taking the special case of simultaneous measurements ($t_1 = t_2$), we obtain the operator rule

$$a^{(m)} = \sum_{l=1}^{m-1} f_l(a^{(l)}) \mapsto \mathsf{A}^{(m)} = \sum_{l=1}^{m-1} f_l(\mathsf{A}^{(l)}). \tag{27}$$

Case 3. $f(a, b) = f_1(a) f_2(b)$

We again consider three sub-cases. First, if measurements **A** and **B** are represented by commuting operators and, if the measurements are sub-system measurements and are performed on the same sub-system, then, in the quantum experiment, they are performed on the same copy of the system. Since f is, therefore, simple, the AVCP applies, so that the relation

$$\begin{aligned} \sum_k c_k P''_k &= \sum_{i,j} f_1(a_i) f_2(b_j) P_i P'_{j|i} \\ &= \sum_i f_1(a_i) f_2(b_i) P_i \end{aligned} \tag{28}$$

holds for all initial states, v_0, of the system. Hence, the generalized operator relation

$$\langle \mathsf{C} \rangle_{t_2} = \big\langle f_1(\mathsf{A}) f_2(\mathsf{B}) \big\rangle_{t_1} \tag{29}$$

holds for all v_0. In the case where measurements **A**, **B**, and **C** are simultaneous, we obtain the operator relation

$$\mathsf{C} = f_1(\mathsf{A}) f_2(\mathsf{B}). \tag{30}$$

Second, if measurements **A** and **B** are represented by commuting operators and are sub-system measurements performed on different sub-systems of a composite system, then they can be performed on the same copy of the system in the quantum experiment, in which case we obtain the same result as above.

Third, if measurements **A** and **B** are represented by non-commuting operators, then the function f is not simple, and the AVCP does not apply.

We can combine the three foregoing sub-cases to obtain the *generalized product rule*:

$$\begin{aligned} c(t_2) &= f_1\big(a(t_1)\big) f_2\big(b(t_1)\big) \\ &\mapsto \langle \mathsf{C} \rangle_{t_2} = \big\langle f_1(\mathsf{A}) f_2(\mathsf{B}) \big\rangle_{t_1} \quad \forall \mathsf{v}_0 \quad \text{if } [\mathsf{A}, \mathsf{B}] = 0. \end{aligned} \tag{31}$$

In the special case where $t_1 = t_2$, we obtain the *product rule*:

$$c = f_1(a) f_2(b) \mapsto \mathsf{C} = f_1(\mathsf{A}) f_2(\mathsf{B}) \quad \text{if } [\mathsf{A}, \mathsf{B}] = 0. \tag{32}$$

1. Some comments on inconsistencies

As mentioned in Sec. II A 2, if the average-value condition is weakened to allow a measurement of AB to be

represented by an operator in the case where $[\mathsf{A},\mathsf{B}]\neq 0$, one is lead to a rule that is often stated, namely

$$f_1(a)f_2(b)\mapsto \frac{1}{2}\big(f_1(\mathsf{A})f_2(\mathsf{B})+f_1(\mathsf{A})f_2(\mathsf{B})\big) \tag{33}$$

However, using this rule, one finds that inconsistencies quickly arise. For example, one can first apply this rule to find that the operator representing a measurement of AB is

$$\widehat{AB}=(\mathsf{AB}+\mathsf{BA})/2, \tag{34}$$

where the notation $\widehat{X}$ is used to denote the operator that represents a measurement of X. One can then apply the rule a second time to find the operator that represents a measurement of A^2B. By treating this measurement as a measurement of $A(AB)$, or as a measurement of $(A^2)B$, one obtains, respectively, either

$$\begin{aligned}\widehat{A(AB)}&=\frac{1}{2}(\mathsf{A}\widehat{AB}+\widehat{AB}\mathsf{A})\\&=\frac{1}{4}\left(\mathsf{A}(\mathsf{AB}+\mathsf{BA})+(\mathsf{AB}+\mathsf{BA})\mathsf{A}\right)\\&=\frac{1}{4}\left(\mathsf{A}^2\mathsf{B}+2\mathsf{ABA}+\mathsf{BA}^2\right),\end{aligned} \tag{35}$$

or

$$\begin{aligned}\widehat{(A^2)B}&=\frac{1}{2}(\widehat{A}^2\mathsf{B}+\mathsf{B}\widehat{A}^2)\\&=\frac{1}{2}(\mathsf{A}^2\mathsf{B}+\mathsf{BA}^2),\end{aligned} \tag{36}$$

which are, in general, inequivalent. Hence, the average-value condition cannot be applied to non-simple functions of observables, even in weakened form, without leading to inconsistencies.

We also remark that, given the AVCP, it cannot consistently be maintained that every classically-described measurement is represented by a quantum measurement, since, under this assumption, the function and sum rules can be applied to a measurement of $(A+B)^2$, with $[\mathsf{A},\mathsf{B}]\neq 0$, to derive Eq. (34) as follows. First, defining $d=a+b$, we use the sum rule to find $\mathsf{D}=\mathsf{A}+\mathsf{B}$, and then use the function rule to find that

$$\begin{aligned}\widehat{D^2}&=\mathsf{D}^2\\&=\mathsf{A}^2+\mathsf{AB}+\mathsf{BA}+\mathsf{B}^2.\end{aligned} \tag{37}$$

Second, since, by assumption, a measurement of AB is represented by a quantum measurement, we can use the sum rule directly to find a measurement of $D^2=A^2+2AB+B^2$:

$$\begin{aligned}\widehat{D^2}&=\widehat{A^2}+2\widehat{AB}+\widehat{B^2}\\&=\mathsf{A}^2+2\widehat{AB}+\mathsf{B}^2.\end{aligned} \tag{38}$$

Equating these expressions for $\widehat{D^2}$, we obtain Eq. (34), which, as we have seen, leads to an inconsistency. If the AVCP is accepted as valid, this inconsistency can only be avoided if we conclude that a measurement of AB (when $[\mathsf{A},\mathsf{B}]\neq 0$) cannot be represented by a quantum measurement, in which case the sum rule cannot be applied to obtain Eq. (38).

In summary, given that measurements of A and B are represented by quantum measurements $\mathbf{A}$ and $\mathbf{B}$, one can use the AVCP to find quantum measurements that represent measurements of $f(A)$, $f_1(A)+f_2(B)$ and, for $[\mathsf{A},\mathsf{B}]=0$, of $f_1(A)f_2(B)$; and, more generally, one can find quantum measurements that represent measurements of $f(A,B)$ when f is simple. The AVCP also implies that a measurement of AB when $[\mathsf{A},\mathsf{B}]\neq 0$ cannot be represented by a quantum measurement. However, this does not appear to be a significant restriction since such measurements do not occur in the Hamiltonian for a system of particles, in either non-relativistic or relativistic physics, in which fundamental forces alone are acting.

B. Temporal Evolution

In this section, we will use the AVCP, together with Postulate 3.4 (see Paper I), to derive the explicit form of the temporal evolution operator for a system in a time-dependent background.

Temporal evolution of the system is represented by a unitary transformation. Specifically, over the course of the interval $[t,t+\Delta t]$, the state $\mathsf{v}(t)$ evolves as

$$\mathsf{v}(t+\Delta t)=\mathsf{U}_t(\Delta t)\mathsf{v}(t), \tag{39}$$

where $\mathsf{U}_t(\Delta t)$ is the unitary matrix that represents temporal evolution of the system during $[t,t+\Delta t]$.

Suppose now that the background of the system is time-independent during this interval. Then we shall write $\mathsf{U}_t(\Delta t)$ as $\mathsf{V}_t(\Delta t)$. Now, for $0\leq\Delta t_1+\Delta t_2\leq\Delta t$, and $\Delta t_1,\Delta t_2$ both positive, we have

$$\mathsf{V}_t(\Delta t_1+\Delta t_2)=\mathsf{V}_{t+\Delta t_1}(\Delta t_2)\mathsf{V}_t(\Delta t_1). \tag{40}$$

But the time-independence of the background implies that $\mathsf{V}_{t+\Delta t_1}(\Delta t_2)=\mathsf{V}_t(\Delta t_2)$. Therefore,

$$\mathsf{V}_t(\Delta t_1+\Delta t_2)=\mathsf{V}_t(\Delta t_2)\mathsf{V}_t(\Delta t_1), \tag{41}$$

which can be solved to yield

$$\mathsf{V}_t(\Delta t_1)=\exp(-i\mathsf{K}_t\Delta t_1), \tag{42}$$

where K_t is a Hermitian matrix.

To determine the nature of K_t, we proceed as follows. In the classical model of a physical system in a time-independent interval $[t,t+\Delta t]$, the classical Hamiltonian for the system is not explicitly dependent upon time during this interval. On the assumption that the classical Hamiltonian is a simple function of the observables of the system, and that the measurements of these observables are represented by quantum measurements [32], it

follows from the AVCP that the corresponding Hamiltonian operator is also not explicitly dependent upon time during this interval, so that, in particular,

$$\mathsf{H}_t = \mathsf{H}_{t+\Delta t}, \tag{43}$$

where H_t denotes the Hamiltonian operator at time t. In addition, in the classical model, the total energy of the system is constant during this interval for all states of the system. Therefore, by the generalized function rule, the relation

$$\langle \mathsf{H}_t \rangle_t = \langle \mathsf{H}_{t+\Delta t} \rangle_{t+\Delta t} \tag{44}$$

holds for any state v. Hence, from Eqs. (43) and (44), it follows that

$$\langle \mathsf{H}_t \rangle_t = \langle \mathsf{H}_t \rangle_{t+\Delta t} \tag{45}$$

for all v. But

$$\langle \mathsf{H}_t \rangle_{t+\Delta t} = \langle \mathsf{H}_t \rangle_t + i\langle [\mathsf{K}_t, \mathsf{H}_t] \rangle_t \Delta t + \mathrm{O}(\Delta t^2). \tag{46}$$

Therefore, the commutator $[\mathsf{K}_t, \mathsf{H}_t] = 0$, which implies that there exist N mutually orthogonal eigenvectors, $\mathsf{v}_1, \ldots, \mathsf{v}_N$, which K_t and H_t share in common. In particular, the state v_j $(j = 1, 2, \ldots, N)$ is an eigenvector of H_t, with some eigenvalue E_j.

Now, if the system is in an eigenstate, v_j, with eigenvalue k_j, of K_t, at time t, the state evolves as

$$\mathsf{v}(t + \Delta t) = e^{-ik_j \Delta t}\, \mathsf{v}(t). \tag{47}$$

Therefore, during the interval $[t, t + \Delta t]$, this state remains an eigenstate of H_t, and is therefore a state of constant energy, E_j, during this interval. In addition, since evolution only affects the overall phase of the state, the observable degrees of freedom of the state are time-independent during this interval. But, by Postulate 3.4, the state $\mathsf{v}(t)$, representing a system in a time-independent background of definite energy, E_j, whose observable degrees of freedom are time-independent, evolves as

$$\mathsf{v}(t + \Delta t) = e^{-iE_j \Delta t/\alpha}\, \mathsf{v}(t). \tag{48}$$

By comparison of Eqs. (47) and (48), we find that $k_j = E_j/\alpha$ holds for $j = 1, 2, \ldots, N$, which implies that $\mathsf{K}_t = \mathsf{H}_t/\alpha$. Hence, in a time-independent background, any state v evolves as

$$\mathsf{v}(t + \Delta t) = \exp\left(-i\mathsf{H}_t \Delta t/\alpha\right) \mathsf{v}(t). \tag{49}$$

In order to generalize to the case of temporal evolution in a time-dependent background, we split the interval $[t, t+\Delta t]$ into intervals of duration ϵ, approximate the evolution during each of these intervals assuming that the background is time-independent, and then take the limit as $\epsilon \to 0$:

$$\begin{aligned} \mathsf{U}_t(\Delta t) &= \mathsf{U}_{t+\Delta t-\epsilon}(\epsilon) \ldots \mathsf{U}_{t+\epsilon}(\epsilon) \mathsf{U}_t(\epsilon) \\ &= \lim_{\epsilon \to 0} \big\{ \mathsf{V}_{t+\Delta t-\epsilon}(\epsilon) \ldots \mathsf{V}_{t+\epsilon}(\epsilon) \mathsf{V}_t(\epsilon) \big\}, \end{aligned} \tag{50}$$

which, upon expansion, yields

$$\mathsf{U}_t(\Delta t) = I - \frac{i}{\alpha} \mathsf{H}_t \Delta t + \mathrm{O}(\Delta t^2), \tag{51}$$

so that

$$i\alpha \frac{d\mathsf{v}(t)}{dt} = \mathsf{H}_t \mathsf{v}(t). \tag{52}$$

The value of the constant α will be determined at the end of Sec. IV A.

IV. REPRESENTATION OF MEASUREMENTS AND SYMMETRY TRANSFORMATIONS

In this section, we shall use the AVCP to obtain (a) the commutation relations $[\mathsf{x}, \mathsf{p}_x] = i\hbar$ and $[\mathsf{L}_x, \mathsf{L}_y] = i\hbar \mathsf{L}_z$ (and cyclic permutations thereof), and a restricted form of Dirac's Poisson bracket rule, (b) the explicit form of the displacement and rotation operators, and (c) the relation between the momentum and displacement operators, and between the angular momentum and rotation operators.

From this point onwards, we shall take the infinite-dimensional form of the abstract quantum formalism as a given.

A. Position, momentum, and displacement operators

We shall proceed in five steps. First, by considering a particular system (a photon moving along the x-axis), we shall obtain the commutation relationship, $[\mathsf{x}, \mathsf{p}_x] = i\alpha$. Second, we shall obtain the explicit co-ordinate representation of the displacement operator $\mathsf{D}_x = \frac{1}{i}\frac{d}{dx}$. Third, we shall obtain the relationship $\mathsf{D}_x = \mathsf{p}_x/\alpha$ that holds between the displacement operator, D_x, and p_x, and thereby obtain the co-ordinate representation of p_x. Fourth, we shall show that the relations obtained are generally valid. Finally, we shall make the identification $\alpha = \hbar$.

1. The position–momentum commutation relationships

Consider a photon moving in the $+x$-direction, where measurements of the x-component of position, the x-component of momentum, and the energy, are represented by the operators $\mathsf{x}, \mathsf{p}_x, \mathsf{H}$, respectively.

First, to determine the relationship of H to the operators x and p_x, we make use of the fact that the relation $H = cp_x$ holds for all classical states (x, p_x) of the system, so that, from the function rule, it follows that $\mathsf{H} = c\mathsf{p}_x$.

Next, to obtain a relation between x and p_x, we make use of the fact that, in the quantum model, the expected

value of x at time $t+\delta t$ can be calculated in two separate ways. First, from the definition of $\langle x\rangle_t$, the relation

$$\begin{aligned}\langle \mathsf{x}\rangle_{t+\delta t} &= \left\langle \mathsf{U}_t^\dagger(\delta t)\,\mathsf{x}\,\mathsf{U}_t(\delta t)\right\rangle_t \\ &= \left\langle \left(1+\frac{i\mathsf{H}\delta t}{\alpha}\right)\mathsf{x}\left(1-\frac{i\mathsf{H}\delta t}{\alpha}\right)\right\rangle_t + O(\delta t^2) \\ &= \langle \mathsf{x}\rangle_t + \frac{i}{\alpha}\delta t\langle \mathsf{H}\mathsf{x}-\mathsf{x}\mathsf{H}\rangle_t + O(\delta t^2) \\ &= \langle \mathsf{x}\rangle_t - \frac{ic}{\alpha}\delta t\left\langle[\mathsf{x},\mathsf{p}_x]\right\rangle_t + O(\delta t^2)\end{aligned} \tag{53}$$

holds for all states, v, of the system. Second, using the generalized function rule, it follows from the classical relation $x(t+\delta t) = x(t) + c\delta t + O(\delta t^2)$ that the relation

$$\langle \mathsf{x}\rangle_{t+\delta t} = \langle \mathsf{x}\rangle_t + c\delta t + O(\delta t^2) \tag{54}$$

holds for all v.

Equating the above two expressions for $\langle x\rangle_{t+\delta t}$, we obtain

$$\mathsf{v}^\dagger\,[\mathsf{x},\mathsf{p}_x]\,\mathsf{v} = i\alpha \tag{55}$$

for all v, which implies that

$$[\mathsf{x},\mathsf{p}_x] = i\alpha. \tag{56}$$

Although a particular system has been used to obtain this commutation relation, we shall later argue that it is generally valid.

2. *Co-ordinate representation of the displacement operator.*

Suppose that, in frame S, the system is in state $\psi(x)$. The probability density function over x' in the frame S', which is displaced a distance $-\epsilon$ along the x-axis, can be calculated in two equivalent ways, according to whether the transformation from frame S to S' is treated as a passive or active transformation. Accordingly, the probability density function over x' can be obtained by performing measurements of x' in frame S' upon the system in state $\psi(x)$, or by performing measurements of x in frame S upon the system in the transformed state, $\exp(-i\epsilon\mathsf{D}_x)\psi(x)$, and substituting x for x' in the resulting probability density function over x.

First, in frame S', let us calculate the probability density function over x' directly. In this frame, the operator x' represents a measurement of x'. In the classical model of the system, the relation

$$x' = x + \epsilon \tag{57}$$

holds for all states (x, p_x) of the system. Hence, by the function rule, we obtain the operator relation

$$\mathsf{x}' = \mathsf{x} + \epsilon. \tag{58}$$

Hence, an eigenstate of x with eigenvalue x is an eigenstate of x' with eigenvalue $x' = x + \epsilon$. Therefore, if a measurement of x on a system in state $\psi(x)$ yields values in the interval $[x, x+\Delta x]$ with probability $|\psi(x)|^2\Delta x$, then a measurement of x' on a system in the same state yields values in the interval $[x', x'+\Delta x']$ with probability density

$$\Pr\left(x'|S',\psi(x)\right) = \left|\psi(x'-\epsilon)\right|^2. \tag{59}$$

Second, in frame S, measurement x is performed on the system in the transformed state $\exp(-i\epsilon\mathsf{D}_x)\psi(x)$, so that the probability density function over x is

$$\Pr\left(x|S,\exp(-i\epsilon\mathsf{D}_x)\psi(x)\right) = \left|\exp(-i\epsilon\mathsf{D}_x)\psi(x)\right|^2. \tag{60}$$

The probability density functions over x' in Eq. (59) and over x in Eq. (60), must agree under the correspondence $x \leftrightarrow x'$. Hence

$$\psi(x-\epsilon) = e^{i\phi(x)}\exp(-i\epsilon\mathsf{D}_x)\psi(x), \tag{61}$$

with $\phi(x)$ being an arbitrary real-valued function of x, which is satisfied for any ϵ if and only if $\phi(x) = 0$ and

$$\mathsf{D}_x = \frac{1}{i}\frac{d}{dx}. \tag{62}$$

3. *The displacement–momentum operator relation.*

In a classical model, the state of a particle subject to measurements of x-position and the x-component of momentum, is given by (x_0, p_{x0}) in some frame of reference, S. Consider the following two experiments.

In the first experiment, measurements of the x components of position and momentum of the particle are made in a reference frame, S', that is displaced by a distance ϵ along the $-x$ axis, resulting in the state (x', p'_x) of the particle relative to the co-ordinates of frame S'. According to the classical model,

$$\begin{aligned}x' &= x_0 + \epsilon \\ p'_x &= p_{x0}.\end{aligned} \tag{63}$$

In the second experiment, the particle is displaced a distance ϵ in the $+x$-direction, and measurements of position and momentum are then performed in frame S, giving the state, (x, p_x), of the particle in frame S as $(x_0+\epsilon, p_{x0})$.

In classical physics, for all states of the particle, the state (x', p'_x), determined by measurements in frame S' upon the undisplaced particle, is numerically identical to the state (x, p_x), determined by measurements in frame S upon the displaced particle. That is,

$$(x', p'_x) = (x, p_x) \tag{64}$$

for all states, (x_0, p_{x0}), of the particle.

Now consider a quantum model of the particle subject to measurements of x and p_x, and let the state of the particle be given by v_0 in frame S. Consider the first experiment. From Eqs. (63), it follows from the generalized

function rule that, in the quantum model of the particle, the relations

$$\begin{aligned}\langle \mathsf{x}' \rangle &= \mathsf{v}_0^\dagger \mathsf{x} \mathsf{v}_0 + \epsilon \\ \langle \mathsf{p}_x' \rangle &= \mathsf{v}_0^\dagger \mathsf{p}_x \mathsf{v}_0,\end{aligned} \tag{65}$$

hold for all quantum states, v_0, of the system.

In the second experiment, the displacement of the particle is a continuous, symmetry transformation of the system, and therefore can be represented by a unitary transformation of the state, v_0, and, in particular, by the operator $\exp(-i\epsilon \mathsf{D}_x)$, where D_x is a Hermitian operator. To first order in ϵ, measurements of x and p_x performed on this state have expected values

$$\begin{aligned}\langle \mathsf{x} \rangle &= \mathsf{v}_0^\dagger (1 + i\epsilon \mathsf{D}_x) \mathsf{x} (1 - i\epsilon \mathsf{D}_x) \mathsf{v}_0 \\ \langle \mathsf{p}_x \rangle &= \mathsf{v}_0^\dagger (1 + i\epsilon \mathsf{D}_x) \mathsf{p}_x (1 - i\epsilon \mathsf{D}_x) \mathsf{v}_0.\end{aligned} \tag{66}$$

From Eq. (64), by the generalized function rule, the average values $\langle x' \rangle$ and $\langle p_x' \rangle$ of Eqs. (65) are respectively equal to the average values $\langle x \rangle$ and $\langle p_x \rangle$ of Eqs. (66) for all v_0. Hence, we obtain that the relations

$$\begin{aligned}\mathsf{v}_0^\dagger (1 + i\epsilon \mathsf{D}_x) \mathsf{x} (1 - i\epsilon \mathsf{D}_x) \mathsf{v}_0 &= \mathsf{v}_0^\dagger \mathsf{x} \mathsf{v}_0 + \epsilon \\ \mathsf{v}_0^\dagger (1 + i\epsilon \mathsf{D}_x) \mathsf{p}_x (1 - i\epsilon \mathsf{D}_x) \mathsf{v}_0 &= \mathsf{v}_0^\dagger \mathsf{p}_x \mathsf{v}_0,\end{aligned} \tag{67}$$

hold for all v_0 to first order in ϵ, which yield the commutation relations

$$\begin{aligned}[\mathsf{x}, \mathsf{D}_x] &= i \\ [\mathsf{p}_x, \mathsf{D}_x] &= 0\end{aligned} \tag{68}$$

From Eqs. (56) and (68), it follows that

$$\begin{aligned}[\mathsf{x}, (\mathsf{D}_x - \mathsf{p}_x/\alpha)] &= 0 \\ [\mathsf{D}_x, (\mathsf{D}_x - \mathsf{p}_x/\alpha)] &= 0.\end{aligned} \tag{69}$$

Now, in the co-ordinate representation, the operators x and D_x are given by x and $-i\alpha\, d/dx$, respectively, and one can readily show that $\{x, -i\alpha\, d/dx\}$ forms an irreducible set [33]. By Schur's lemma [34], it therefore follows from Eqs. (69) that

$$\mathsf{D}_x = \frac{\mathsf{p}_x}{\alpha} + \gamma \mathsf{I}, \tag{70}$$

where γ is real since the operator $(\mathsf{D}_x - \mathsf{p}_x/\alpha)$ is Hermitian. For a given displacement, ϵ, the constant γ results in the same overall shift of phase of any state, v, of a system, and therefore produces no physically observable effects on the system. Hence, γ can be set equal to zero without any loss of generality, so that we obtain

$$\mathsf{D}_x = \frac{\mathsf{p}_x}{\alpha}. \tag{71}$$

Analogous relationships for the displacement operator corresponding to displacements in the y and z directions can be obtained in a similar way.

Finally, from Eqs. (62) and (71), we find

$$\mathsf{p}_x = \frac{\alpha}{i} \frac{d}{dx}. \tag{72}$$

4. Generality.

The representations of x- and p_x-measurements have been obtained above by considering, in the first step, a quantum model of a particular physical system, namely a photon moving along the x-axis. In the general case of a photon moving in an arbitrary direction, measurements of x, y, z, and p_x, p_y, p_z are sub-system measurements, and can therefore be represented in the model of the composite system consisting of a photon, subject to measurements chosen from a measurement set generated by a measurement of $\mathbf{r} = (x, y, z)$, by the operators x, y, z and $-i\alpha\, \partial/\partial x, -i\alpha\, \partial/\partial y, -i\alpha\, \partial/\partial z$, respectively.

These representations of measurements of position and momentum are also more generally valid for other systems, as we shall explain below.

a. State-determined measurements. Suppose that, in the classical framework, a measurement of A is performed on a system, and the outcome is determined by the state of the system alone. That is, in particular, the outcome is independent of the background of the system or of any parameters (such as charge or rest mass) that describe intrinsic properties of the system. We shall then say that this measurement is a *state-determined* measurement. For example, the outcome of a position measurement on a particle is determined by the state of the particle, and is independent of whether or not the particle is in an electromagnetic field and is independent of the mass or charge of the particle. In general, any measurement of an observable that is a function only of the degrees of freedom of the state of the system is a state-determined measurement. In contrast, a measurement of the total energy of a system is, in general, dependent upon not only the state of the system, but also upon the background of the system, and is therefore not a state-determined measurement.

Now, consider two quantum models of two different physical systems, system 1 and system 2, in different backgrounds, with respect to the measurement set generated by measurements $\mathbf{A}_1$ and $\mathbf{A}_2$, respectively, where $\mathbf{A}_1$ and $\mathbf{A}_2$ represent a measurement of A performed on the respective systems. Suppose, further, that the two models have the same dimension. If the measurement of A is state-determined when performed on both systems 1 and 2, then, by the AVCP, it follows that $\langle \mathsf{A}_1 \rangle = \langle \mathsf{A}_2 \rangle$ holds for all states, v, where operators $\mathsf{A}_1, \mathsf{A}_2$ represent measurements $\mathbf{A}_1, \mathbf{A}_2$, respectively. It follows at once that the operators $\mathsf{A}_1, \mathsf{A}_2$ are identical.

Hence, provided that two systems admit classical models with respect to a measurement of A that is state-determined, and admit quantum models of the same dimension with respect to measurements $\mathbf{A}_1$ and $\mathbf{A}_2$, the operators that represent $\mathbf{A}_1$ and $\mathbf{A}_2$ in the respective models must be identical.

Therefore, if state-determined measurements of x and p_x are performed on any system, then, in a quantum model of the system subject to measurements in

the measurement set containing quantum measurements that represent measurements of x and p_x, where these measurements yield a continuum of possible outcomes, their representations are the same as those obtained above. Therefore, the commutation relations involving x, p_x and D_x are also generally valid. Similar conclusions hold for measurements of y, z and p_y, p_z.

Therefore, in the case of a particle where the interaction energy in the Hamiltonian is obtained from a scalar potential that is dependent on position only, in which case the measurements of position and momentum are state-determined, the above representations are valid. Below, we shall consider a physically important case where the measurement of momentum is not state-determined.

b. Particle in a magnetic field. In the case of a charged particle in a magnetic field background described in the Hamiltonian framework, the state of the particle is $(\mathbf{x}, \dot{\mathbf{x}})$, but the generalized co-ordinates are taken to be $(\mathbf{x}, \mathbf{p})$, where $\mathbf{p} = m\dot{\mathbf{x}} + e\mathbf{A}$, where m and e are the mass and charge of the particle, respectively, and $\mathbf{A} = (A_x, A_y, A_z)$ is the vector potential. In this case, $\mathbf{p}$ depends both upon the state of the particle and the state of the background. Therefore, a measurement of $\mathbf{p}$ is not a state-determined measurement, and the foregoing argument cannot be used to argue that the operators representing the measurements of the components, p_x, p_y, p_z, of $\mathbf{p}$ are those derived above. Instead, we reason as follows.

First, for a particle with state $(\mathbf{x}, \dot{\mathbf{x}})$ in a magnetic field, in the argument of Sec. IV A 1, the commutation relation for the x-component of the motion in Eq. (56) becomes

$$\left[\mathsf{x}, m\dot{\mathsf{x}}\right] = i\alpha, \tag{73}$$

Then, from $\mathbf{p} = m\dot{\mathbf{x}} + e\mathbf{A}$, the sum rule gives

$$\mathsf{p}_x = m\dot{\mathsf{x}} + eA_x(\mathsf{x}, \mathsf{y}, \mathsf{z}), \tag{74}$$

which, together with Eq. (73), implies that

$$[\mathsf{x}, \mathsf{p}_x] = i\alpha, \tag{75}$$

as before.

Second, we note that, in the classical framework, the momentum $\mathbf{p}$ as defined above is invariant under displacement of the reference frame. Therefore, Eqs. (68) remain unchanged, and, using Eq. (75), we obtain $\mathsf{D}_x = \mathsf{p}_x/\alpha$. Third, and finally, the argument leading to the coordinate representation of D_x remains unchanged since the argument only involves measurements of position, which are state-determined measurements. Therefore, the explicit representation of p_x remains that given in Eq. (72), and similarly for the y- and z-components of the motion.

5. *Identification of $\alpha = \hbar$.*

At this point, having obtained explicit representations for position and momentum measurements, it is possible to use the operator rules to write down the explicit Schroedinger equation for a structureless electron in a hydrogen atom. By solution of the equation, and by comparing the energy levels of the electron either with those found in Bohr's model or with those found by experiment, one can establish that the constant α is equal to $\hbar$.

6. *Remark on applications.*

The formal rules derived above allow the quantum theoretic modeling of a non-relativistic particle in an arbitrary classical background consisting of gravitational and electromagnetic fields, which leads to the non-relativistic Schroedinger equation. In the case of a multi-particle system, the rules (not discussed here) for dealing with identical particles are, additionally, required.

In addition, the above rules allow the modeling of a photon without consideration of polarization degrees of freedom (leading to a complex wave equation), a structureless relativistic particle (leading to the Klein-Gordon equation), and a relativistic particle with internal degrees of freedom (which, with the appropriate auxiliary assumptions, leads to the Dirac equation).

B. Angular momentum and rotation operators

We shall proceed in three steps. First, we shall obtain the commutation relation $[\mathsf{L}_x, \mathsf{L}_y] = i\hbar\mathsf{L}_z$, and cyclic permutations thereof, up to an additive constant. Second, we shall obtain the commutator relations that hold between the rotation operators, $\mathsf{R}_x, \mathsf{R}_y, \mathsf{R}_z$, and the angular momentum operators, and shall then use these relations to determine the value of the additive constant. Third, we shall determine the relations that hold between the rotation and angular momentum operators, and indicate how the explicit representations of the angular momentum operators and rotation operators can be determined.

1. *Components of Angular Momentum.*

Consider an experimental set-up where, in the classical model of the set-up, measurements are performed upon a classical spin, with magnetic moment $\boldsymbol{\mu}$, which determine the values of the rectilinear components of angular momentum of the system. Suppose that the measurements of the components of angular momentum along the x-, y-, and z-directions, and the measurement of energy, are represented by the operators $\mathsf{L}_x, \mathsf{L}_y, \mathsf{L}_z$, and H, respectively.

In particular, consider a set-up where a magnetic field, $\mathbf{B}$, is applied to the spin. In the classical model of this set-up, the energy associated with the interaction is $-\boldsymbol{\mu} \cdot \mathbf{B}$. Since $\boldsymbol{\mu} = q\mathbf{L}/2m$, where q and m are the charge and mass, respectively, of the spin, and $\mathbf{L} =$

(L_x, L_y, L_z) is its angular momentum vector, the energy can be written as $-(q/2m)\,\mathbf{B}\cdot\mathbf{L}$. By the sum rule (Sec. III A), the quantum mechanical Hamiltonian is given by

$$\mathsf{H} = -\frac{q}{2m}(B_x\mathsf{L}_x + B_y\mathsf{L}_y + B_z\mathsf{L}_z), \tag{76}$$

where B_x, B_y and B_z are the rectilinear components of $\mathbf{B}$.

The application of a magnetic field to a classical spin causes its angular momentum vector, $\mathbf{L}$, to rotate about the axis along which the magnetic field is applied by an angle that is proportional both to $|\mathbf{B}|$ and to the duration for which the field is applied. Let the rotation matrix corresponding to a rotation about axis a be denoted $R_a(\theta)$, where θ is the angle of rotation. From the properties of rotation matrices, it follows that

$$R_x(\epsilon)R_y(\epsilon) - R_y(\epsilon)R_x(\epsilon) = R_z(\epsilon^2) - I, \tag{77}$$

where ϵ is an infinitesimal angle, and $R_a(\epsilon)$ is an infinitesimal rotation which can be implemented by application of a magnetic field $\mathbf{B}$ along the axis a for some time δt. Using this relationship, it is possible to deduce the commutation relations that hold between the quantum mechanical operators, $\mathsf{L}_x, \mathsf{L}_y, \mathsf{L}_z$, in the following way.

The unitary evolution corresponding to the application of a magnetic field $\mathbf{B}$ to a spin for a time δt is

$$\mathsf{U}(\delta t) = \exp\left(-\frac{i}{\hbar}\mathsf{H}\delta t\right). \tag{78}$$

If magnetic fields of equal strength are applied for equal times, δt, along the x, y, and z-axes, respectively, the corresponding unitary evolution is given to first order in δt, respectively, by

$$\begin{aligned}\mathsf{U}_1(\delta t) &= 1 - \frac{i\epsilon}{\hbar}\mathsf{L}_x\\ \mathsf{U}_2(\delta t) &= 1 - \frac{i\epsilon}{\hbar}\mathsf{L}_y\\ \mathsf{U}_3(\delta t) &= 1 - \frac{i\epsilon}{\hbar}\mathsf{L}_z.\end{aligned} \tag{79}$$

Define $\mathrm{proj}(\mathsf{v})$ as the operation upon the quantum state, v, of a spin which returns a three-dimensional vector, $\langle\mathsf{L}\rangle$, with components $\langle\mathsf{L}_x\rangle, \langle\mathsf{L}_y\rangle$ and $\langle\mathsf{L}_z\rangle$.

If the application of a magnetic field, $\mathbf{B} = B_z\mathbf{k}$, say, to a classical spin causes a rotation of $\mathbf{L}$ by angle θ, then, by the generalized operator rule in Eq. (27), in the quantum model of the spin, the application of the field rotates the vector $\langle\mathsf{L}\rangle = \mathrm{proj}(\mathsf{v})$ by the angle θ about the z-axis. From Eq. (77), it therefore follows that, for any v,

$$\mathrm{proj}(\mathsf{U}_1(\epsilon)\mathsf{U}_2(\epsilon)\mathsf{v}) - \mathrm{proj}(\mathsf{U}_2(\epsilon)\mathsf{U}_1(\epsilon)\mathsf{v}) = \mathrm{proj}(\mathsf{U}_3(\epsilon^2)\mathsf{v}) - \mathrm{proj}(\mathsf{v}). \tag{80}$$

Using the definitions

$$\begin{aligned}\mathsf{v}_1 &= \mathsf{U}_3(\epsilon^2)\mathsf{v} = \mathsf{v} + \delta\mathsf{v}_1\\ \mathsf{v}_2 &= \mathsf{U}_1(\epsilon)\mathsf{U}_2(\epsilon)\mathsf{v} = \mathsf{v} + \delta\mathsf{v}_2\\ \mathsf{v}_3 &= \mathsf{U}_2(\epsilon)\mathsf{U}_1(\epsilon)\mathsf{v} = \mathsf{v} + \delta\mathsf{v}_3,\end{aligned}$$

where

$$\begin{aligned}\delta\mathsf{v}_1 &= -\frac{i}{\hbar}\epsilon^2\mathsf{L}_z\mathsf{v}\\ \delta\mathsf{v}_2 &= -\left[-\frac{i}{\hbar}\epsilon(\mathsf{L}_x + \mathsf{L}_y) - \frac{1}{\hbar^2}\epsilon^2\mathsf{L}_x\mathsf{L}_y\right]\mathsf{v}\\ \delta\mathsf{v}_3 &= -\left[-\frac{i}{\hbar}\epsilon(\mathsf{L}_x + \mathsf{L}_y) - \frac{1}{\hbar^2}\epsilon^2\mathsf{L}_y\mathsf{L}_z\right]\mathsf{v},\end{aligned}$$

equation (80) becomes

$$\mathrm{proj}(\mathsf{v} + \delta\mathsf{v}_2) - \mathrm{proj}(\mathsf{v} + \delta\mathsf{v}_3) = \mathrm{proj}(\mathsf{v} + \delta\mathsf{v}_1) - \mathrm{proj}(\mathsf{v}). \tag{81}$$

Equating the x-components of this equation, we obtain

$$\mathsf{v}^\dagger\mathsf{L}_x(\delta\mathsf{v}_2 - \delta\mathsf{v}_3) + (\delta\mathsf{v}_2 - \delta\mathsf{v}_3)^\dagger\mathsf{L}_x\mathsf{v} = \mathsf{v}^\dagger\mathsf{L}_x\delta\mathsf{v}_1 + \delta\mathsf{v}_1^\dagger\mathsf{L}_x\mathsf{v}, \tag{82}$$

and, inserting the explicit forms of the $\delta\mathsf{v}_i$, we obtain the commutation relation

$$\left[\mathsf{L}_x,\ \mathsf{L}_z + \frac{i}{\hbar}[\mathsf{L}_x, \mathsf{L}_y]\right] = 0. \tag{83a}$$

Equating the y- and z-components similarly, one obtains the relations

$$\left[\mathsf{L}_y,\ \mathsf{L}_z + \frac{i}{\hbar}[\mathsf{L}_x, \mathsf{L}_y]\right] = 0 \tag{83b}$$

$$\left[\mathsf{L}_z,\ \mathsf{L}_z + \frac{i}{\hbar}[\mathsf{L}_x, \mathsf{L}_y]\right] = 0 \tag{83c}$$

By inspection, the above commutation relations have the solution

$$\left[\mathsf{L}_x, \mathsf{L}_y\right] = i\hbar\mathsf{L}_z + i\gamma_1\mathsf{I}, \tag{84a}$$

where γ_1 is real constant since the operators $i[\mathsf{L}_x, \mathsf{L}_y]$ and L_z are hermitian. We shall later show that this solution is, in fact, the most general one.

The discussion leading to this result can be repeated to yield the relations

$$\left[\mathsf{L}_y, \mathsf{L}_z\right] = i\hbar\mathsf{L}_x + i\gamma_2\mathsf{I} \tag{84b}$$

$$\left[\mathsf{L}_z, \mathsf{L}_x\right] = i\hbar\mathsf{L}_y + i\gamma_3\mathsf{I}. \tag{84c}$$

In order to determine the values of γ-factors, we require the commutation relations between the rotation operators, $\mathsf{R}_x, \mathsf{R}_y, \mathsf{R}_z$, and $\mathsf{L}_x, \mathsf{L}_y, \mathsf{L}_z$, which we shall now derive.

2. Rotation–angular momentum commutation relations.

Let an infinitesimal clockwise rotation of a frame of reference by angle ϵ about the z-axis be represented by unitary transformation $\exp(-i\epsilon\mathsf{R}_z)$, where R_z is Hermitian.

Now consider a set-up where measurements of L_x, L_y, and L_z are performed on a system in the original and

in the transformed frame of reference. In the classical model of this situation, the outcomes of the measurements performed in the original (unprimed) and transformed (primed) frames are, to first order in ϵ, related by

$$\begin{pmatrix} L'_x \\ L'_y \\ L'_z \end{pmatrix} = \begin{pmatrix} 1 & -\epsilon & 0 \\ \epsilon & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} L_x \\ L_y \\ L_z \end{pmatrix}. \tag{85}$$

By the generalized operator rule in Eq. (27), it follows that, in the quantum model of the situation, the relation

$$\begin{pmatrix} \langle \mathsf{L}'_x \rangle \\ \langle \mathsf{L}'_y \rangle \\ \langle \mathsf{L}'_z \rangle \end{pmatrix} = \begin{pmatrix} 1 & -\epsilon & 0 \\ \epsilon & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \langle \mathsf{L}_x \rangle \\ \langle \mathsf{L}_y \rangle \\ \langle \mathsf{L}_z \rangle \end{pmatrix} \tag{86}$$

holds for all states, v, of the system. Using the relation

$$\langle \mathsf{L}'_x \rangle = \mathsf{v}^\dagger (1 + i\epsilon \mathsf{R}_z) \mathsf{L}_x (1 - i\epsilon \mathsf{R}_z) \mathsf{v} + O(\epsilon^2), \tag{87}$$

we find from Eq. (86) that

$$\langle \mathsf{L}_x \rangle + i\epsilon \langle [\mathsf{R}_z, \mathsf{L}_x] \rangle = \langle \mathsf{L}_x \rangle - \epsilon \langle \mathsf{L}_y \rangle \tag{88}$$

for all v, which implies that

$$\big[\mathsf{R}_z, \mathsf{L}_x\big] = i\mathsf{L}_y. \tag{89a}$$

Proceeding similarly for $\langle \mathsf{L}'_y \rangle$ and $\langle \mathsf{L}'_z \rangle$, we obtain

$$[\mathsf{R}_z, \mathsf{L}_y] = -i\mathsf{L}_x \tag{89b}$$

$$[\mathsf{R}_z, \mathsf{L}_z] = 0. \tag{89c}$$

The commutation relations for R_x and R_y can be obtained by parallel arguments.

3. Angular momentum commutation relations.

Left-multiplying Eq. (84c) by R_z, and applying Eqs. (89a)–(89c), we obtain

$$\big[\mathsf{L}_z, \mathsf{L}_x\big]\mathsf{R}_z - i\big[\mathsf{L}_y, \mathsf{L}_z\big] = (i\hbar\mathsf{L}_y + i\gamma_3 \mathsf{I})\,\mathsf{R}_z + \hbar\mathsf{L}_x. \tag{90}$$

Using the Eq. (84c) right-multiplied by R_z, this implies that

$$\big[\mathsf{L}_y, \mathsf{L}_z\big] = i\hbar\mathsf{L}_x. \tag{91a}$$

Parallel arguments applied to Eqs. (84a) and (84b) using the commutation relations between $\mathsf{R}_x, \mathsf{R}_y$ and $\mathsf{L}_x, \mathsf{L}_y, \mathsf{L}_z$ yield

$$\big[\mathsf{L}_z, \mathsf{L}_x\big] = i\hbar\mathsf{L}_y \tag{91b}$$

$$\big[\mathsf{L}_x, \mathsf{L}_y\big] = i\hbar\mathsf{L}_z. \tag{91c}$$

4. Explicit form of angular momentum operators

From the classical relation $L^2 = L_x^2 + L_y^2 + L_z^2$, it follows from the sum rule that $\mathsf{L}^2 = \mathsf{L}_x^2 + \mathsf{L}_y^2 + \mathsf{L}_z^2$. Using this relation and the above commutation relations for $\mathsf{L}_x, \mathsf{L}_y$ and L_z, explicit representations of these operators for finite N can be obtained and the irreducibility of the representations of $\mathsf{L}_x, \mathsf{L}_y, \mathsf{L}_z$ can be shown [35]. Therefore, by Schur's lemma, the solution given in Eq. (84a) is the most general solution of Eqs. (83a)–(83c), and similarly for the solutions given in Eqs. (84b) and (84c).

Although the representations of $\mathsf{L}_x, \mathsf{L}_y$ and L_z have been obtained by considering a particular physical system, they are generally valid on account of the argument given in Sec. IV A 4. Therefore the commutation relations for $\mathsf{L}_x, \mathsf{L}_y$ and L_z are generally valid.

5. Rotation–angular momentum relations, and explicit form of the rotation operators.

Using the commutation relationships for $\mathsf{L}_x, \mathsf{L}_y$ and L_z, it follows from Eqs. (89a)–(89c) that $(\hbar\mathsf{R}_z - \mathsf{L}_z)$ commutes with $\mathsf{L}_x, \mathsf{L}_y$, and L_z. Since $\{\mathsf{L}_x, \mathsf{L}_y, \mathsf{L}_z\}$ is an irreducible set, it follows from Schur's lemma that

$$\hbar\mathsf{R}_z - \mathsf{L}_z = \gamma \mathsf{I}, \tag{92}$$

where γ is a real-valued constant since R_z and L_z are Hermitian. For any given ϵ, a non-zero value of γ results in the same overall change of phase for all states transformed by $\exp(-i\epsilon\mathsf{R}_z)$, and so cannot give rise to observable consequences. Hence, γ can be set to zero without loss of generality. Therefore, $\mathsf{R}_z = \mathsf{L}_z/\hbar$, and, similarly, $\mathsf{R}_x = \mathsf{L}_x/\hbar$ and $\mathsf{R}_y = \mathsf{L}_y/\hbar$.

Using the explicit representations of $\mathsf{L}_x, \mathsf{L}_y, \mathsf{L}_z$ for any given dimension N, the explicit representation of the rotation operators follows at once from these rotation–angular momentum relations. The explicit co-ordinate representations of the rotation operators in the infinite-dimensional case can also be determined by an argument similar to that used earlier to determine the explicit form of the displacement operators.

C. Commutators and Poisson brackets

In this section, we shall obtain a relation between the Poisson Bracket, $\{A, B\}$, and the commutator $[\mathsf{A}, \mathsf{B}]$, where A and B are the classical observables of a physical system describable in the classical Hamiltonian framework, and A, B are the operators that represent measurements of these observables. Dirac's Poisson Bracket rule asserts the relation

$$[\mathsf{A}, \mathsf{B}] = i\hbar\widehat{\{A, B\}}, \tag{93}$$

where $\widehat{\{A, B\}}$ is the operator that represents a measurement of $\{A, B\}$. Below, we shall derive this relation using the AVCP in the case where B is the Hamiltonian.

Consider the Hamiltonian model of a system with state $(q_1,\ldots,q_N;p_1,\ldots,p_N)$ where $N \geq 1$. The temporal rate of change of the function $F(q_1,\ldots,q_N;p_1,\ldots,p_N)$ is given in terms of the Hamiltonian, $H(q_1,\ldots,q_N;p_1,\ldots,p_N)$, by

$$\begin{aligned}\dot{F} &= \{F,H\} \\ &= \sum_{i=1}^{N}\left\{\frac{\partial F}{\partial q_i}\frac{\partial H}{\partial p_i} - \frac{\partial H}{\partial q_i}\frac{\partial F}{\partial p_i}\right\}. \end{aligned} \tag{94}$$

Consider the quantum model of the system with state v, where the measurements of the q_i and the p_i are represented by operators q_i and p_i, respectively. If H is simple, then, by the AVCP, a measurement of H can be represented by the operator H; otherwise, according to the AVCP, it is not possible to describe the temporal evolution of the system in the quantum model. If both of the functions $\dot{F}$ and $\{F,H\}$ are simple, then, by the AVCP, they are represented by the operators $\widehat{\dot{F}}$ and $\widehat{\{F,H\}}$, respectively, and from Eq. (94), by the generalized function rule, the relation

$$\left\langle \widehat{\dot{F}} \right\rangle_t = \left\langle \widehat{\{F,H\}} \right\rangle_t, \tag{95}$$

holds for all v.

Now, in the classical model, the function $\dot{F}$ is defined, for all states, as

$$\dot{F} = \lim_{\Delta t\to 0}\left\{\frac{F(t+\Delta t)-F(t)}{\Delta t}\right\}. \tag{96}$$

If $F(t)$ and $F(t+\Delta t)$ are both simple, then, according to this definition, $\dot{F}(t)$ is also simple, and, using the generalized sum rule (regarding the measurement of $F(t+\Delta t)$ as the one being implemented in terms of measurements of $F(t)$ and $\dot{F}(t)$), we obtain the relation

$$\left\langle \widehat{\dot{F}} \right\rangle_t = \lim_{\Delta t\to 0}\frac{1}{\Delta t}\left\{\langle \mathsf{F}\rangle_{t+\Delta t} - \langle \mathsf{F}\rangle_t\right\}, \tag{97}$$

which holds for all v, with the operator F representing a measurement of F.

If the functions $\{F,H\}$ and $F(t)$ are both simple, then, since $F(t+\Delta t) = F(t) + \{F,H\}\Delta t$, it follows that $F(t+\Delta t)$ is also simple. In that case, both Eqs. (95) and (97) hold for all v. Equating these two expressions for $\langle\widehat{\dot{F}}\rangle_t$, we obtain the relation

$$\begin{aligned}\left\langle \widehat{\{F,H\}} \right\rangle_t &= \lim_{\Delta t\to 0}\frac{1}{\Delta t}\left\{\left\langle\left(1+\frac{i}{\hbar}\mathsf{H}\Delta t\right)\mathsf{F}\left(1-\frac{i}{\hbar}\mathsf{H}\Delta t\right)\right\rangle_t \right. \\ &\qquad \left. - \langle \mathsf{F}\rangle_t\right\} \\ &= i\hbar^{-1}\left\langle [\mathsf{H},\mathsf{F}]\right\rangle_t, \end{aligned} \tag{98}$$

which holds for all v, so that

$$i\hbar\widehat{\{F,H\}} = [\mathsf{F},\mathsf{H}]. \tag{99}$$

Hence, we obtain Eq. (93) in the special case where $\mathsf{B} = \mathsf{H}$, subject to the condition that the functions A, B and $\{A,B\}$ are simple.

Using this relationship, we can readily evaluate useful commutation relationships. For example, for a photon with state (x,p_x), setting $F = x$, and $H = cp_x$ (the Hamiltonian for a photon moving along the $+x$-direction), we find $\{F,H\} = c$. Hence, since the functions F, H, and $\{F,H\}$ are simple, Eq. (99) immediately gives $[\mathsf{x},\mathsf{p}_x] = i\hbar$.

If one or more of the functions F, H, and $\{F,H\}$ is not simple, then Eq. (99) does not follow from the above argument. To take a specific example, suppose that, for a system with state (x,p_x), where $[\mathsf{x},\mathsf{p}_x] = i\hbar$, we choose $F = x^3$ and $H = \gamma p_x^3$, where γ is a constant. We can then apply the function rule to obtain the corresponding operators $\mathsf{F} = \mathsf{x}^3$ and $\mathsf{H} = \gamma\mathsf{p}_x^3$, and use these to find

$$[\mathsf{F},\mathsf{H}] = 3i\gamma\hbar\left(\mathsf{x}^2\mathsf{p}_x^2 + \mathsf{x}\mathsf{p}_x^2\mathsf{x} + \mathsf{p}_x^2\mathsf{x}^2\right). \tag{100}$$

However, the function $\{F,H\} = 9\gamma x^2p_x^2$ is not simple, which implies that the AVCP cannot be used to write down an operator which represents a measurement of $\{F,H\}$. If we were nonetheless to apply the Hermitization rule in Eq. (33) to a measurement of $\{F,H\}$ (in spite of the inconsistencies to which we have shown this would lead) we would obtain

$$i\hbar\widehat{\{F,H\}} = \frac{9i\gamma\hbar}{2}\left(\mathsf{x}^2\mathsf{p}_x^2 + \mathsf{p}_x^2\mathsf{x}^2\right), \tag{101}$$

but this differs from $[\mathsf{F},\mathsf{H}]$ by the constant $2\gamma\hbar^3$. Since the expected value of $\gamma\hbar\mathsf{x}^2\mathsf{p}_x^2$ is itself of order $\gamma\hbar^3$, there is no guarantee that the difference between Eqs. (100) and (101) will be negligible.

The question of whether Eq. (93) holds in the more general case where the function B cannot be treated as the classical Hamiltonian of the system is not discussed here.

V. ARBITRARINESS IN THE FUNCTIONS $f(\chi_i)$ AND $\tilde{f}(\chi_i)$

In Paper I, we found that the functions $f(\chi_i)$ and $\tilde{f}(\chi_i)$ (not to be confused the function f that forms part of the AVCP) are $\pm\cos(a\chi_i+b)$ and $\pm\sin(a\chi_i+b)$, respectively, where $a,b\in\mathbb{R}$ and $a\neq 0$. Having obtained the temporal evolution operator and commutation relationships, we are now able to show that the choices made in Paper I of the positive signs for f and $\tilde{f}$ and of $a=1$ and $b=0$ do not involve a loss of generality.

Let us first consider the case where positive signs are chosen for both f and $\tilde{f}$. With constants a and b left in place, the derivation given in Paper I is altered as follows. First, with $Q_{a|i} = \cos(a\chi_i+b)$ and $Q_{b|i} = \sin(a\chi_i+b)$,

and defining $\tilde{\chi}_i = a\chi_i + b$, we obtain

$$\begin{aligned}\mathbf{Q} &= (\sqrt{P_1}Q_{a|1}, \sqrt{P_1}Q_{b|1}, , \ldots, \sqrt{P_N}Q_{b|N}) \\ &= (\sqrt{P_1}\cos\tilde{\chi}_1, \sqrt{P_1}\sin\tilde{\chi}_1, \ldots, \sqrt{P_N}\sin\tilde{\chi}_N),\end{aligned} \tag{102}$$

where the P_i are the probabilities of the observed outcomes of the measurement, $\mathbf{A}$, with respect to which $\mathbf{Q}$ is written.

Second, the invariance condition (Postulate 3.2) requires that there is no change in the probabilities of the observable outcomes of measurement $\mathbf{A}$ performed on an evolved state if an arbitrary constant $\chi_0 \in \mathbb{R}$ is added to each of the χ_i in the initial state. Hence, in terms of the $\tilde{\chi}_i$, an arbitrary constant $\tilde{\chi}_0$ may be added to each of the $\tilde{\chi}_i$. Therefore, from the argument given in Paper I, the state can, with respect to measurement $\mathbf{A}$, be represented as the complex vector

$$\mathsf{v} = \begin{pmatrix} \sqrt{P_1}e^{i\tilde{\chi}_1} \\ \sqrt{P_2}e^{i\tilde{\chi}_2} \\ \vdots \\ \sqrt{P_N}e^{i\tilde{\chi}_N} \end{pmatrix}, \tag{103}$$

with physical transformations being represented as unitary or antiunitary transformations as found previously.

Third, Postulate 5 implies that $\tilde{\chi}_{ij} = \tilde{\chi}_i^{(1)} + \tilde{\chi}_j^{(2)}$, which implies that the composite systems rule remains $\mathsf{v} = \mathsf{v}^{(1)} \otimes \mathsf{v}^{(2)}$, where $\mathsf{v}^{(1)}, \mathsf{v}^{(2)}$ are the states of the subsystems, and v is the state of the composite system.

Fourth, by inspection of the argument used above to find the temporal evolution operator, one finds that $\mathsf{U}_t(\Delta t) = \exp(-ia\mathsf{H}\Delta t/\alpha)$. In addition, the foregoing arguments lead to $\mathsf{p}_x = -i(a/\alpha)\, d/dx$. If one writes down the Schroedinger equation implied by these relations, one obtains $\alpha/a = \hbar$, which implies that $\mathsf{U}_t(\Delta t) = \exp(-i\mathsf{H}\Delta t/\hbar)$ as before.

In summary, irrespective of the values of a, b, one obtains the same abstract quantum formalism. Furthermore, although the constants a, b appears as a connection between χ_i and $Q_{a|i}$, we shall now show that they are unimportant insofar as observable predictions of the formalism, namely the prediction of the values and probabilities of the observable outcomes of measurements, are concerned. To do so, we shall consider an experiment where a system is prepared using measurement $\mathbf{A}$ and, after undergoing a physical transformation (either active or passive), is subject to measurement $\mathbf{A}'$.

We shall first establish that the explicit form of a measurement operator that represents any given measurement is independent of a and b. First, consider the example of the operators x and p_x. The explicit form of these operators is determined by (a) the commutation relation $[\mathsf{x}, \mathsf{p}_x] = i\hbar$, (b) the measurement–transformation relation $\mathsf{D}_x = \mathsf{p}_x/\hbar$, and (c) the relation $\mathsf{D}_x = -i\, d/dx$. In this case, one finds by inspection of the above derivation that these relations are all independent of a and b. More generally, the explicit form of measurement operators is determined by (a) commutation relationships between measurement operators, (b) the relation between measurement operators and transformation operators, and (c) the explicit form of transformation operators. First, since the Dirac Poisson bracket rule derived in Sec. IV C is independent of a and b, the commutation relations between measurement operators are independent of a and b. Second, the measurement–transformation relations and the explicit form of the transformation relations are directly obtained via the AVCP from the classical relation between the outcomes of measurements performed in the original and transformed frames of reference, and are therefore independent of a and b (as illustrated by the above arguments leading to $\mathsf{D}_x = \mathsf{p}_x/\hbar$ and $\mathsf{D}_x = -i\, d/dx$). Therefore, in general, measurement operators are independent of a and b.

Next, we consider the operators that represent physical transformations. First consider active transformations. The general temporal evolution operator, $\mathsf{U}_t(\Delta t) = \exp(-i\mathsf{H}_t\Delta t/\hbar)$ is not explicitly dependent upon a and b. The operator H_t is a function of the measurement operators (such as x, p_x for a particle moving along the x-axis) relevant to the system, and is directly obtained (via the AVCP) from the classical relation between the classical Hamiltonian, H, and the classical observables that determine the state of the system, and is therefore also independent of a and b. Since the measurement operators are themselves independent of a and b, it follows that the operator representing any given temporal evolution is independent of a and b.

Second, the explicit form of operators representing passive transformations (such as displacement or rotation of a reference frame) can be obtained by directly transposing the relevant classical relations (which connect the co-ordinates in the original and transformed frames) into the quantum framework via the AVCP (as illustrated by the derivation of the co-ordinate representation of D_x above), and can be readily seen to therefore also be independent of a and b. Hence, a unitary or antiunitary operator that represents any given physical transformation is independent of a and b.

Now consider the above-mentioned experiment where measurement $\mathbf{A}$ is performed and, say, outcome 1 is obtained. With respect to measurement $\mathbf{A}$, the resulting state is $\mathsf{v} = (1, 0 \ldots, 0)^{\mathsf{T}}$ up to an irrelevant overall phase. Suppose the system now undergoes a physical transformation (either active or passive), and a second measurement, $\mathbf{A}$, is performed. Since the physical transformation is represented by a unitary or antiunitary operator that is independent of a and b as noted above, the outcome probabilities of the second measurement are unaffected by the value of a and b. This example includes the general case where the second measurement is measurement $\mathbf{A}'$ since the latter can, by Postulate 1.2 (see Paper I), be represented by an arrangement consisting of measurement $\mathbf{A}$ immediately preceded and followed by suitable temporal evolution of the system. Therefore, in the most general experiment we are considering, the outcome probabilities of the observable outcomes are also independent of a

and b.

Finally, we note that, since the explicit form of the measurement operators is independent of a and b, their eigenvalues are also independent of a and b. Therefore, in summary, we find that both the probabilities and the values of the observable outcomes of measurements in the above general experiment are independent of a and b. Therefore, the values of a and b can, without loss of generality, conveniently be chosen to be $a = 1$ and $b = 0$.

If one chooses the signs of f and $\tilde{f}$ not to be both positive, then this is equivalent to choosing positive signs for f and $\tilde{f}$ but changing the values of a and b to some other values, a' and b', respectively. Specifically, if one chooses the signs $(+,-)$, then $a' = -a$ and $b' = -b + \pi$; if $(-,+)$, then $a' = -a$ and $b' = -b$; and, if $(-,-)$, then $a' = a$ and $b' = b + \pi$. But we have already shown that the choice of the constants a and b is unimportant, and hence the change in their values is unimportant. Therefore, the signs can, without loss of generality, both be chosen to be positive.

VI. DISCUSSION

The derivation presented in this paper has shown that, using a clearly-motivated physical principle, it is possible to derive the explicit form of the temporal evolution operator given the postulates of Paper I, and to derive the formal rules of quantum theory in a systematic manner from appropriately chosen relations known to hold in classical physics. The derivation provides several physical insights into the formal rules.

The first insight is that the classical description of a measurement (such as 'a measurement of x^2') leaves open the possibility of more than one implementation, and that, when modeled in the quantum framework, these implementations are, in general, not equivalent.

Second, it is possible to impose a simple average-value condition that must be satisfied by an operator that can be said to represent a classical implementation of a measurement. This condition implies that many implementations cannot be represented by an operator, and can therefore be eliminated from consideration. That is, one finds that there are implementations which, although acceptable in the classical framework, cannot be represented by operators in the quantum framework without violating a very mild average-value condition.

Third, in the case of an implementation that satisfies the average-value condition, the operator that represents the implementation is uniquely determined by the average-value condition provided that the function, f, that describes the implementation, is simple. One also finds that those implementations that satisfy the average-value condition are are represented by the same operator, so that it is possible to represent a measurement of f by a unique operator. If f is not simple, then it does not appear to be possible to apply the average-value condition, even in a weakened form, without inconsistencies arising.

Fourth, we have found that the AVCP is incompatible with the assumption that every classically-described measurement on a system is represented by a quantum measurement in the quantum model of the system. For example, the AVCP implies that a measurement of AB does not have an operator representation if $[\mathsf{A}, \mathsf{B}] \neq 0$.

The fifth insight rests on the fact that, rather surprisingly, the AVCP enables formal rules of each of the four types described in the Introduction (operator rules, commutation relations, transformation operators, and measurement–transformation relations) to be obtained in a uniform manner. Consequently, one can see that the difference between these types of rules depends simply upon whether the classical relations that one is taking over into the quantum framework are relations between measurements performed at the same time (leading to the operator rules), at different times (leading to commutation relations for measurement operators), or in different frames of reference (leading to measurement–transformation relations and to the explicit forms of transformation operators). In short, from the perspective provided by the derivation, the commutation relation $[\mathsf{x}, \mathsf{p}_x] = i\hbar$ is no more elusive in its origin than the operator relation $\mathsf{H} = \mathsf{p}_x^2/2m$.

Sixth, the derivation provides a clearer physical foundation to many particular formal rules that are commonly used in quantum theory. For example, the commutation relation $[\mathsf{L}_x, \mathsf{L}_y] = i\hbar\mathsf{L}_z$ is ordinarily derived in the infinite-dimensional quantum formalism for a particle (by transposing the classical relation $\mathsf{L}_z = xp_y - yp_x$, and cyclic permutations thereof, into the quantum framework using the operator rules), and is then assumed, without further justification, to also hold in the finite-dimensional case. Here, we have obtained this commutation relation directly for finite- and infinite-dimensional quantum systems, and have done so in a manner that makes clear its connection with the properties of rotations. Similarly, a restricted form of Dirac's Poisson bracket rule has been derived in a systematic manner using the AVCP without making use of abstract analogies.

Finally, we remark that the general notion of average-value correspondence is already familiar in elementary quantum mechanics through Ehrenfest's theorem [22], which shows that the motion of a particle modeled in the quantum framework is, on average, approximated by the behavior of the particle when described classically. However, the possibility that such a correspondence might serve as the basis for a constructive principle that allows the formal rules of quantum theory to be determined by appropriately-chosen classical relations does not appear to have been widely explored [36]. It has been shown here that it is possible to formulate the notion of average-value correspondence in the form of an exact physical principle which, roughly speaking, allows the logic of Ehrenfest's argument to be reversed, enabling the often physically obscure formal rules of quantum theory to be derived in a systematic manner from familiar relations known to hold in classical physics.

Acknowledgments

I am indebted to Steve Gull and Mike Payne for their constant support and invaluable comments, and to Tetsuo Amaya and Matthew Donald for discussions and invaluable comments.

[1] P. Goyal. An information-theoretic approach to quantum theory, I: The abstract quantum formalism. 2007.

[2] W. K. Wootters. *The acquisition of information from quantum measurements*. PhD thesis, University of Texas at Austin, 1980.

[3] C. Rovelli. Relational quantum mechanics. *Int. J. Theor. Phys.*, 35:1637–1678, 1996. See also quant-ph/9609002v2.

[4] S. Popescu and D. Rohrlich. Causality and nonlocality as axioms for quantum mechanics. 1997. See quant-ph/9709026v2.

[5] A. Caticha. Consistency, amplitudes and probabilities in quantum theory. *Phys. Rev. A*, 57:1572, 1998.

[6] A. Caticha. Insufficient reason and entropy in quantum theory. 1999. See quant-ph/9810074v2.

[7] J. Summhammer. Maximum predictive power and the superposition principle. 1999. See quant-ph/9910039.

[8] Č. Brukner and A. Zeilinger. Operationally invariant information in quantum measurements. *Phys. Rev. Lett.*, 83:3354–3357, 1999.

[9] Č. Brukner and A. Zeilinger. Conceptual inadequacy of the Shannon information in quantum mechanics. *Phys. Rev. A*, 63(022113), 2001.

[10] Č. Brukner and A. Zeilinger. Information and fundamental elements of the structure of quantum theory. 2002. See quant-ph/0212084v1.

[11] R. Clifton, J. Bub, and H. Halvorson. Characterising quantum theory in terms of information-theoretic constraints. 2003. See quant-ph/0211089v2.

[12] A. Grinbaum. Elements of information-theoretic derivation of the formalism of quantum theory. *Int. J. Quant. Inf.*, 1(3):289–300, 2003. See also quant-ph/0306079.

[13] A. Grinbaum. *The Significance of Information in Quantum Theory*. PhD thesis, Ecole Polytechnique, Paris, 2004. See quant-ph/0410071.

[14] B. R. Frieden and Soffer. Lagrangians of physics and the game of Fisher-information transfer. *Phys. Rev. E*, 52:2274–2286, 1995.

[15] M.J.W. Hall and M. Reginatto. Schroedinger equation from an exact uncertainty principle. *J. Phys. A*, 35(14):3289–3303, 2002.

[16] M. Reginatto. Derivations of the equations of nonrelativistic mechanics using the principle of minimum Fisher information. *Phys. Rev. A*, 58:1775–8, 1998.

[17] J. von Neumann. *Mathematical Foundations of Quantum Mechanics*. Princeton University Press, 1955.

[18] H. J. Groenewold. On the principles of elementary quantum mechanics. *Physica*, 12:405–60, 1946.

[19] T. F. Jordan. Why $-i\nabla$ is the momentum. *Am. J. Phys.*, 43(12), 1975.

[20] T. F. Jordan. *Linear Operators for Quantum Mechanics*. John Wiley and Sons, Inc., 1969.

[21] L. E. Ballentine. *Quantum Mechanics : A Modern Development*. World Scientific, 1999.

[22] P. Ehrenfest. Bemerkung über die angenäherte gültigkeit der klassischen mechanik innerhalb der quantenmechanik. *Z. Phys.*, 45(7–8):455–457, 1927.

[23] P.A.M. Dirac. *Principles of Quantum Mechanics*. Oxford Science Publications, fourth edition, 1999.

[24] David Bohm. *Quantum Theory*. Dover Publications, 1989.

[25] C. Isham. *Lectures on Quantum Theory*. Imperial College Press, 1995.

[26] J. F. Cornwell. *Group Theory in Physics: An Introduction*. Academic Press, 1997.

[27] The basic operator rules of quantum theory are (i) Rule 1 (Function rule): If a measurement of A is represented by operator A, then a measurement of $f(A)$ is represented by $f(\mathsf{A})$; (ii) Rule 2 (Sum rule): If measurements of A and B are represented by operators A, B, respectively, then a measurement of $f_1(A) + f_2(B)$ is represented by $f_1(\mathsf{A}) + f_2(\mathsf{B})$ (and similarly for more than two observables), and (iii) Rule 3 (Product rule): If measurements of A and B are represented by commuting operators A, B, respectively, then a measurement of $f_1(A)f_2(B)$ is represented by $f_1(\mathsf{A})f_2(\mathsf{B})$ (and similarly for more than two observables). A more general rule that is often employed is: (iv) Rule $3'$ (Hermitization rule): If measurements of A and B are represented by operators A, B, respectively, then a measurement of $f_1(A)f_2(B)$ is represented by $(f_1(\mathsf{A})f_2(\mathsf{B}) + f_2(\mathsf{B})f_1(\mathsf{A}))/2$. The above rules are generally assumed to apply to arbitrary Hermitian operators.

[28] See [23] (Sec. 11) and [17] (Sec. IV.1), for example.

[29] See, for example, [17] (Sec. IV.1), [24] (Sec. 9.12—9.15), and [25] (Sec. 5.2.1). We shall discuss one such example in Sec. III A 1.

[30] A measurement that is a degenerate form of a measurement $\mathbf{A}$ is defined operationally in Paper I. Such a measurement has $N' < N$ possible outcomes, and can be represented as an N-dimensional degenerate Hermitian operator (with N' distinct eigenvalues).

[31] A sub-system measurement is defined in Paper I as a measurement performed on a single sub-system of a composite system.

[32] If the measurements of the observables of which the classical Hamiltonian is a function are not represented by quantum measurements, then it is not possible to write down the operator corresponding to the classical Hamiltonian. Similarly, if the classical Hamiltonian is *not* a simple function of the observables of the system, then the form of the operator that represents a measurement of energy cannot be determined by the AVCP and, therefore, the explicit form of $\mathsf{U}_t(\Delta t)$ cannot be obtained from the argument given in the text. However, these assumptions do not appear to represent a significant restriction in any fundamental cases of interest (in both non-relativistic and relativistic cases).

[33] See [21], Appendix 2.

[34] See, for example, Ref. [21], Appendix 1.

[35] See [26] (Ch. 10, Sec. 3), for example.

[36] Refs. [17, 18, 24] mention the general idea of average-value correspondence in their discussion of the operator rules of quantum theory. For example, Groenewold [18] (Eqs. (1.32)–(1.34)) remarks that the sum rule is equivalent to a condition on the expectations of the respective operators, but the idea is not formulated in a manner that is sufficiently systematic to derive the operator rules, and no attempt is made to derive the any of the other types of formal rule (such as the commutation relations) using average-value correspondence. Bohm [24] clearly articulates the idea that average-value correspondence can be used as a constraint on quantum theory, and uses it to determine particular instances of the function rule (Secs. 9.5–9.21) and to determine the Hamiltonian operator that represents a non-relativistic particle (Secs. 9.24–9.26). However, the idea is not systematically formulated and applied beyond these special cases.